PCT WORLD INTELLECTUAL PROPI
International Bi

WO 9606625A1

INTERNATIONAL APPLICATION PUBLISHED UNDER

(51) International Patent Classification 6 :
A61K 35/16, 39/00, 39/395

A1

(11) International Publication Number: WO 96/06625

(43) International Publication Date: 7 March 1996 (07.03.96)

(21) International Application Number: PCT/US95/10791

(22) International Filing Date: 25 August 1995 (25.08.95)

(30) Priority Data:
08/296,625 26 August 1994 (26.08.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/296,625 (CON)
Filed on 26 August 1994 (26.08.94)

(71) Applicant *(for all designated States except US)*: ELI LILLY AND COMPANY [US/US]; Lilly Corporate Center, Indianapolis, IN 46285 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ILL, Charles, R. [US/US]; 1098 Oceanic Drive, Encinitas, CA 92024 (US). LUDWIG, James, Richard [US/US]; 8548 Tidewater West, Indianapolis, IN 46236 (US). RATHNACHALAM, Radhakrishnan [IN/US]; 3793 Lattice Court, Carmel, IN 46032 (US).

(74) Agent: LAMMERT, Steven, R.; Barnes & Thornburg, 1313 Merchants Bank Building, 11 South Meridian Street, Indianapolis, IN 46204 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.

(54) Title: ANTIBODY CONSTRUCTS WITH CDR SWITCHED VARIABLE REGIONS

(57) Abstract

CDR grafted recombinant antibodies are provided which have at least one CDR switched variable domain wherein one or more of the heavy chain CDRs from one chain of the donor antibody are grafted into the framework regions of the light chain of the acceptor antibody. To enhance the binding of the CDRs as well as the secretion level of multi-chain constructs, the recombinant antibodies are altered using techniques of molecular modeling.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

Title

ANTIBODY CONSTRUCTS WITH CDR SWITCHED VARIABLE REGIONS

Background of the Invention

This invention relates to antigen binding molecular agents useful as *in vitro* diagnostic or *in vivo* imaging and therapy agents. More specifically, the invention relates to the preparation of antibody-derived proteins useful for diagnosing, imaging and therapy of cancer, cardiovascular lesions, infections, and other pathological states. In particular, this invention relates to recombinant antibody-like proteins with reduced immunogenic properties which can be efficiently expressed in eukaryotic cells.

Native antibodies are comprised of four protein chains, two shorter 'light' chains and two longer 'heavy' chains. The chains are associated in a specific three dimensional structure. Each of the four chains consists of a series of linked domain structures.

These domains are structurally related, incorporating a structural unit known as the immunoglobulin fold. Each chain contains one variable domain, encoded by a variable exon, and a number of constant domains, encoded by constant exons, the number being determined by whether the chain is heavy or light and, for heavy chain, determined by

the class of heavy chain. The number of heavy chain constant domains is three for the most commonly occurring class of immunoglobulins, IgG. The constant region of the light chain consists of a single domain, C_L. When the variable domains are properly folded, according to the dictates of the protein sequence, the intact antibody provides a structure with specific binding properties. If the intact antibody molecule is envisioned as a Y-shape, the stem of the Y (Fc) is formed by surface complementarity of the C_{H-2}, hinge, and C_{H-3} portions of the constant regions of the two heavy chains, which extend beyond the light chains. In addition, the two heavy chains are covalently linked through a number of disulfide linkages, the number of disulfide linkages varying between different antibody classes (i.e. IgG, IgM, IgD, IgE, IgA) and subclasses (e.g. IgG_1, IgG_2, IgG_3, IgG_4). The constant region of the gamma-1 heavy chain, for example, includes three constant domains, C_{H-1}, C_{H-2}, and C_{H-3}, with C_{H-1} linked to C_{H-2} by an extended linker region called the hinge. The five classes of antibodies are determined in the main by their differing heavy chains - thus the IgA, IgD, IgE, IgG and IgM classes have alpha, delta, epsilon, gamma and mu type heavy chains, respectively. Each of these types of heavy chain are characterized by having generally conserved amino acid sequences in their constant domains and hinge regions, regardless of the antigen to which they bind. There are additionally two classes of light chains, lambda and kappa, the latter being more abundant in many mammalian species including mouse (ratio of kappa:lambda of 90:10) and human (ratio of kappa:lambda of 60:40). As with the heavy chains, each class of light chain has a generally conserved constant domain sequence regardless of the antigen to which the variable domain of the chain binds.

The variable domains are complementary, so that one heavy and light chain pair joins to form each arm of the antibody. Thus, the amino terminus of each arm contains a region (Fv) containing the antigen binding variable domains of one light and one heavy chain. Each variable domain

contains three complimentarity determining regions (CDRs) characterized by highly variable protein sequences between different antibodies. Each CDR is framed by two of the four framework regions (FRs) present in each variable region, thus creating an alternating sequence of FR-CDR-FR-CDR-FR-CDR-FR-(constant domain).

Antibody specificity and affinity are governed by the sequence and structure of the CDRs. Outside of the CDRs (i.e. within the FRs), the variable domains of the light and heavy chains have the same general structure, albeit with noticeable and functionally significant differences in sequence. The four FRs largely adopt a ß-sheet conformation and are joined by connecting loops which incorporate the CDRs. The CDRs are held in close proximity by the FRs. Note that it is not always necessary to have complementary pair variable domains from one heavy and light chain to obtain binding, as is found in native antibodies. Ward, *et al.*, Nature, 341:544-546 (1989), demonstrated that some V_H domains by themselves have the capability of binding antigens.

Various forms of antibodies and antibody fragments are known for use in delivering drugs and toxins to specific sites within the body. Similarly, radiolabeled antibodies and antibody constructs can be administered *in vivo* for detecting and imaging or treating tumors, thrombi, infection, and other disease states. These immunotherapeutic and imaging agents target a binding site on a particular tissue or cell type, for example, a specific antigen associated with a tumor or thrombus. As a result, other tissues or cells do not accumulate the attached radioisotope, drug or toxin to the same extent. Thus, the risk of toxicity to normal tissue during systemic administration of drugs and radiolabels is considerably lessened, and concomitantly the dose of the therapeutic agent may be lowered.

Another approach in the case of antibodies for therapy and diagnosis is to use antigen binding fragments. Antibody fragments display more rapid specific targeting, less non-specific accumulation in the liver and spleen (due

to the absence of the Fc portion), and a faster rate of clearance from the blood stream than intact antibodies. Due to these characteristics, antibody fragments permit the use of radioisotopes with short half lives, such as ^{99m}Tc, ^{186}Rh, and the like, as well as isotopes with longer half lives such as ^{90}Y and ^{111}In.

The greatest amount of information to date has been obtained with antibody fragments which have been produced by enzymatic digestion of antibodies, with or without chemical reduction. Digestion with papain cleaves the molecule above the hinge region, containing the interchain disulfide bonds linking the two heavy chains. The resultant fragments include two identical F_{AB} fragments, containing the heavy and light chain variable domains, referred to generally by the abbreviations V_H and V_L, respectively, the light chain constant domain, C_L, and the first heavy chain constant domain, C_{H-1}, as well as a small portion of the hinge region. When the intact antibody is digested instead with the proteolytic enzyme pepsin, the cleavage is below the disulfide bonds of the hinge region and results in a bivalent molecule having the F_{AB} regions from both arms linked by the disulfides in a larger segment of the hinge than in the F_{AB}. The resulting fragment is called an $F(ab')_2$ fragment. Upon reduction of the disulfide bonds, the $F(ab')_2$ fragment produces two Fab' fragments. However, the enzymatic cleavage process often results in low yields and a significant loss of binding properties. (See Wahl, *et al.*, J. Nucl. Med., 24:317-325, 1983). Therefore, the search continues for targeting molecules having specificity, enhanced binding activity, minimal non-specific binding, and a shorter half-life in vivo than intact antibodies. This is especially true for in vivo diagnostic (imaging) applications.

While antibody fragments have advantages for many applications, the intact antibody has advantages for many therapeutic approaches. Naked antibody therapy (i.e. therapy utilizing antibody molecules which are not coupled to drugs, radioisotopes, or toxins) often requires effector functions

located in the Fc portion for action. This Fc portion is absent from most fragments. In addition, radioimmunotherapy may be more effective with intact molecules as the total dose delivered is a function of residence time at the tumor, which is uniformly higher for intact antibody molecules over fragments due to the same factors that cause fragments to be more rapidly cleared from the blood stream.

It is also possible to directly express immunoglobulin deletion mutants such as Fab or $F(ab')_2$-like fragments, using recombinant DNA techniques. In one such procedure, an Fd' fragment (i.e. the portion of the immunoglobulin heavy chain found in the $F_{AB'}$ molecule) was expressed in *E. coli* (Cabilly, *et al.*, Proc. Natl. USA 81:3273-3277 (1984)). Ward, *et al.*, in "Binding Activities of a Repertoire of Single Immunoglobulin Variable Domains Secreted from *Escheria coli*", Nature 341:544-546 (1989), also describe expression of isolated heavy chain variable domain genes from *E. coli* to form a type of binding fragment known as a "single domain antibody." In another method described by Gillies, PCT Patent Application No. PCT/US91/00633, specific constant region domains of the human gamma heavy chain, such as the C_{H-2} domain, were eliminated to enhance the binding activity and eliminate effector functions (such as complement activation and Fc receptor binding) of the recombinant molecule over that of the native antibody.

Ideally, human antibodies and antibody fragments would be used for immunotherapy and immunodiagnosis of humans in order to avoid the undesired immune responses often caused by administering non-human immunoglobulins to them. However, human antibodies of appropriate specificity and affinity are difficult to obtain. For instance, conventional hybridoma techniques yield species hybrid cell lines that are frequently unstable and often produce IgM antibodies, instead of the more desirable IgG class of antibodies. An IgM molecule is expressed primarily as a pentamer made up of five identical subunits (IgM monomers), each containing two heavy and two light chains. IgM monomers have, as a rule,

affinities that are too low for therapeutic and imaging applications. Therefore, methods utilizing genetic engineering have been developed for "humanizing" non-human immunoglobulins. In the initial attempts, chimeric antibodies were fashioned by replacing the entire variable domain of human antibodies with those of another species (usually murine). (See Morrison, *et al.*. European Patent Application No. EP 0 173 494 and PCT Patent Application No. PCT/US91/01844.) However, many chimeric antibodies have proven immunogenic because they still contain sufficient non-human protein sequences to generate an immune response.

In a further step towards "humanization," the CDR's of human (acceptor) antibody species have been replaced by those of another (donor) species, so that the framework regions and the constant domains are entirely or predominantly human immunoglobulin, and only the CDR portion of the recombinant antibody is non-human (See European Patent Application Publication No. EP 0 239 400 by Winter, *et al.*). These constructs, commonly known as "CDR-grafted antibodies," can also be made as antibody fragments, (See Winter, *et al.*, European Patent Application Publication No. EP 0 239 400 and Adair, *et al.*, PCT Patent Application No. PCT/GB91/01108) or as single chain antibodies (U.S. Patent No. 4,946,778 (8/7/91) issued to Ladner, *et al.*, and U.S. Patent Nos. 5,132,405 (7/21/92) and 5,091,513 (2/25/92) issued to Huston, *et al.*). However, grafting of the donor CDR regions into the acceptor protein framework can displace the donor binding regions out of their optimal conformations and impair binding affinity in the resultant product. Adair *et al.*, in PCT Patent Application No. PCT/GB90/02017, disclose a method for restoring the CDRs to their native conformation by replacing certain key amino acid residues in the acceptor antibody framework regions to agree with those residues in corresponding regions of the native donor antibody. This procedure increases the binding efficiency of the donor CDRs but at the same time can increase the immunogenicity of the

construct, since non-human residues are introduced into the human part of the construct.

In addition to the above problems in humanizing a non-human antibody, the process of producing vectors containing genomic DNA for encoding humanized antibodies has proven difficult due to the size of these human genes.

Accordingly, there exists in the art a need for more and better genetically engineered antibodies with lowered immunogenicity but with sufficient antigen-binding affinity and specificity to be useful for in vivo detection of disease, for therapy, and for a combination thereof, such as for tumor imaging and cancer therapy. The need also exists for recombinant antibodies that are easily expressed in prokaryotic or eukaryotic host cells in commercially useful quantities and which accumulate in normal tissue in acceptably low amounts. Particularly of interest are recombinant antibodies with reduced immunogenicity (for instance a CDR-grafted antibody, and fragments thereof, comprised of human framework regions and constant domains) that bind quickly to their target sites and have other preferred pharmacokinetic properties.

Since smaller forms of antibodies, such as fragments, are less immunogenic than large intact antibodies, the combination of CDR grafting with small molecular size offers significant advantages for most in vivo applications. However, intact forms also have advantages in applications, such as radioimmunotherapy, where long residence times at the tumor are essential for maximum therapeutic effect.

Another approach to overcoming immunogenicity is the development of multiple reagents having common binding characteristics, but different structures. For example, use of different human frameworks with the same CDRs provides a different overall surface to the host immune system. More directly related to the current invention, use of frameworks from different human immunoglobulin chains provides unique molecular structures, either light chain CDRs with heavy chain frameworks or *vice versa*.

The multiple reagents described above can be used in at least three ways. First, employing different molecular forms in consecutive rounds of therapy can decrease the likelihood of generating an immune response to any one form. Similarly, administering a cocktail combining various forms, decreases the amount of any individual form administered, again decreasing the likelihood of a specific immune response. Finally, alternate molecular forms can be held in reserve, to be administered after an immune response develops to the first form administered.

There exists a need for recombinant antibodies with increased specificity. These higher specificity antibodies should be expressed from mammalian cells in order to have the proper glycosylation, and should be expressed by the cells in practical amounts. In order to impart desirable pharmacokinetic properties, it is further desirable that the recombinant antibodies be fragments of whole antibodies. Finally, it is desirable that these recombinant antibodies be as non-immunogenic as possible. This goal can be accomplished by reducing the size of the construct, by humanizing the construct to the extent possible, and by replacing heavy chain framework regions with light chain framework regions.

Many of the novel molecules embraced by the present invention provide multiple small, humanized forms, which are structurally distinct from native and other recombinant types of humanized antibodies and their fragments, but conserve affinity and specificity.

SUMMARY OF THE INVENTION

The present invention encompasses a recombinant antibody or fragment thereof, and DNA and RNA sequences therefor, comprised of at least one light chain variable domain, which domain, in turn, comprises three CDRs wherein one or more of the CDRs is derived from [identical to or closely resemble(s)] the amino acid sequence of the

corresponding CDR(s) of a heavy chain variable domain of one (donor) antibody and further comprises four framework regions wherein one or more of the amino acid sequence of framework regions are derived from the amino acid sequence of the corresponding framework region(s) from the light chain variable domain of the same or a different (acceptor) antibody, and pharmaceutical compositions containing such antibodies or fragments.

The invention also encompasses DNA sequences encoding such recombinant antibodies or fragments thereof, and vectors containing these DNA sequences in addition to host cells transfected by these vectors.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a schematic representation depicting a recombinant fragment defined herein as a CSV_L fragment. In the example depicted, the CSV_L fragment is fused at the fragment's carboxy terminus to a peptide that chelates metal ions. The illustrative CSV_L fragment also consists of all four framework regions from the V_L domain of an acceptor antibody and all three CDR regions from the V_H domain of a donor antibody.

Figure 2 is a schematic representation depicting a recombinant fragment defined herein as a Heavybody. The Heavybody consists of a CSV_L fragment and a C_L domain.

Figure 3 is a schematic representation depicting a recombinant fragment defined herein as a Kappabody fragment. The Kappabody fragment has two chains: one a Heavybody and the other a CDR-grafted light chain. Preferably, the two chains are connected by a disulfide bond.

Figure 4 is a schematic representation depicting a recombinant molecule defined herein as an Intact Kappabody. This molecule comprises two heavy chains, wherein both of the

heavy chain variable domains have been replaced by CSV_L fragments, and two light chains, wherein both light chains are CDR-grafted.

Figure 5 is a schematic representation depicting a recombinant molecule defined herein as an ScF_V-CSV_L fragment. As the title implies, the Figure depicts a CSV_L fragment bound by a short peptide linker to a CDR-grafted V_L domain.

Figure 6 provides a linear array of the sequences of light chain variable regions of eight antibodies whose atomic coordinates have been deposited in the Brookhaven Protein Data Bank (PDB). The identifiers used in this Figure correspond to PDB file names and antibody names as shown in **Table 2**. The sequences contained within bold boxes represent consensus SCRs. The light boxes associated with SCR5 enclose the SCRs common only to FB4 and each individual sequence of the array. The NSCRs in each sequence are found in the sequence segments outside of (and between, except for NSCR N,1 and NSCR 7,C) the bold boxes. Dots represent gaps introduced into the sequences in order to align the columns in the array.

Figure 7 provides a linear array of the sequences of heavy chain variable regions of eight antibodies from the Brookhaven Data Base. The Brookhaven antibodies are referred to by the identifiers of **Table 2**. The sequences contained within bold boxes represent consensus SCRs. The light boxes associated with SCR1 enclose the SCRs common to only FB4 and each individual sequence of the array. The NSCRs in each sequence are found in the sequence segments outside of (and between, except for NSCR N,1 and NSCR 10,C) the bold boxes. Dots represent gaps introduced into the sequences in order to align the SCRs in the array.

Figure 8 shows the sequence array of the ZCE025 light chain variable region aligned with the Brookhaven

sequences shown in **Figure 6**. The segments of ZCE025 corresponding to the consensus SCRs are contained within bold boxes. Kabat defined CDR residues are in bold. CDR-associated residues are in bold italics.

Figure 9 provides the sequence array of the ZCE025 heavy chain variable region aligned with the Brookhaven sequences shown in **Figure 7**. The segments of ZCE025 corresponding to the consensus SCRs are contained within bold boxes. Kabat-defined CDR residues are in bold. CDR-associated residues are in bold italics.

Figure 10 provides a sequence array in which the sequence of IM9 light chain variable region has been aligned with the Brookhaven sequences shown in **Figure 6**. The IM9 segments corresponding to the consensus SCRs are contained within bold boxes.

Figure 11 depicts a sequence array in which the sequence of IM9 heavy chain variable region has been aligned with the Brookhaven sequences shown in **Figure 7**. The IM9 segments corresponding to the consensus SCRs are contained within bold boxes.

Figure 12 The variable region of the CSVL(HB) containing the light chain variable region of IM9 grafted with the Kabat-defined CDRs from the heavy chain of ZCE025, aligned with the heavy and light chain variable regions of IM9 and ZCE025. Structurally homologous regions between pairs of antibodies are enclosed by boxes.

Figure 13 shows the amino acid sequence of the IM9 light chain variable domain CDR-grafted with CDR's derived from the heavy chain of ZCE025. Lower case letters represent residues from IM9 human V_K domain; upper case letters represent residues from ZCE025 murine V_H domain; @ represents a glycosylation site; * designates CDR-supporting framework

residues from the donor antibody; $ designates residues involved in domain association and ß designates residues that are common to both the V_H domain of ZCE025 and the V_K domain of IM9.

Figure 14 is a restriction map of the 9 Kb BamHI fragment containing the IM9 kappa gene in bacteriophage lambda EMBL3. The MboI termini generated by the partial genomic digest, were reconstructed as BamHI sites. The left and right lambda arms are 20 and 9 Kb, respectively. The exons are represented by solid boxes.

Figure 15 is a restriction map of pBluescript®KS- (commercially available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037) containing IM9 kappa BamHI/BstEII insert from the 5'-end of the IM9 kappa gene subcloned from the 9 Kb BamHI fragment of **Figure 14**. The BstEII site was eliminated by filling in the 5' overhang and cloning into the EcoRV site of pBluescript®KS-. The exons are represented with solid boxes and the Amp^r gene is represented with a box.

Figure 16 is a map showing the primers for overlap PCR mutagenesis of the IM9 kappa gene 5'-end from BamHI to BstEII. The two sets of primers flanking the variable exon specify the addition of Sfi sites on each side of the exon. The location of the MstII site ablation is indicated 5' to the open box representing signal exon I.

Figure 17 is a restriction map of the IM9 kappa expression vector pGIM9kappa. Coding regions are represented by stippled boxes with arrows indicating the direction of transcription. In clockwise order from the ClaI site, the vector consists of the following fragments: a ClaI - BamHI fragment containing the ampicillin resistance gene, the SV 40 promoter, the mycophenolic acid resistance gene, and the SV 40 polyadenylation site; and a BamHI - ClaI fragment

containing the IM9 kappa promoter, the IM9 kappa signal exon, the IM9 kappa signal intron, the IM9 kappa variable exon, the IM9 kappa major intron, including the kappa enhancer, the IM9 kappa constant region exon, and the IM9 kappa polyadenylation site plus 3 Kb of downstream sequence.

Figure 18 shows a restriction map of the pGIM9k/hZCE(CSV_L)-kappa expression vector. In clockwise order 5' to 3' are: the BamHI to SfiI fragment containing the IM9 light chain promoter and signal exon; the SfiI to SfiI fragment containing the CSV_L exon and the 3' end of the major intron; the SfiI to MstII fragment containing the remainder of the major intron (including the IM9 light chain enhancer), the IM9 C_k constant exon, and the IM9 kappa 3' untranslated region; and the MstII to BamHI fragment containing the pSV2gpt (enhancer minus) vector. The solid boxes with arrows indicate open reading frames.

A DETAILED DESCRIPTION OF THE INVENTION

The present invention embraces genetically engineered CDR-grafted recombinant antibodies or antigen-binding fragments comprised of at least one CDR switched light chain variable domain (hereafter referred to as a "CSV_L" fragment or domain), which domain, in turn, comprises three CDRs wherein the amino acid sequence of one or more of the CDRs is derived from the amino acid sequence of the corresponding CDR(s) of a heavy chain variable domain of one (donor) antibody and further comprises four framework regions wherein one or more of the framework regions are derived from the amino acid sequence as the corresponding framework regions(s) from the light chain variable domain of the same or a different (acceptor) antibody. The recombinant antibodies, and the corresponding antigen-binding fragments thereof, will be referred to collectively herein as "CSV_L recombinant antibodies". It will be understood by one skilled in the art that the CSV_L recombinant antibodies can

contain CDRs and FRs from donor and acceptor antibodies of widely divergent origins. Thus, the donor and acceptor antibodies do not have to be from the same species, and whether they are from the same species or not they certainly do not have to be of the same class or subclass. Thus, one could use a murine Ig-alpha donor antibody and a rabbit Ig-gamma acceptor antibody to construct a CSV_L fragment of the instant invention. Similarly, one could use a murine IgG-2a donor antibody and a human IgG-4 acceptor antibody to construct such a fragment.

Five types of CSV_L recombinant antibodies comprise the preferred embodiments of the present invention. The first is the CSV_L fragment itself (see **Figure 1**); the second is a single chain derivative termed a "heavybody" (see **Figure 2**), which is composed of a CSV_L-containing fragment fused through the C-terminus to the N-terminus of a light chain constant domain. A third preferred embodiment is termed a kappabody fragment, which comprises a heavybody chain combined with a CDR-grafted light chain, preferably covalently linked by a disulfide bridge between the two light chain constant domains (see **Figure 3**). The latter light chain differs in general from its CSV_L counterpart in that the CDR-grafted chain has CDRs derived from a donor light chain variable domain substituted for the native CDRs in the acceptor light chain variable domain, versus substitution with donor heavy chain CDRs in the case of a CSV_L domain. A further preferred embodiment is termed an intact kappabody (**see Figure 4**). The intact kappabody resembles an intact CDR-grafted antibody (with all four variable domains having at least one CDR replaced with a non-native CDR of the same type of chain (i.e. heavy or light); differing in that the two CDR-grafted heavy chain variable domains are replaced by two CSV_L domains. The fifth preferred embodiment is termed a single chain chain-switched variable fragment and is defined as a CSV_L domain bonded to a CDR-grafted light chain variable domain throughout a short peptide linker, generally no more than 25 amino acid residues (see **Figure 5**). The symbol used

in the Specification for this embodiment is "ScFv(CSV_L)". The C-terminal end of the CDR-Grafted V_L domain can be fused to the N-terminus of the CSV_L domain through the peptide linker, or vice versa.

As with the CSV_L fragment, one skilled in the art will realize that the Heavybody, the kappabody fragment, the intact kappabody and the ScFv(CSV_L) fragment offer a wide array of choices for donor and acceptor antibodies. Thus, taking the heavybody as an example, the donor antibody could be a murine IgA_1, the Framework Region(s) and the C_L could be from a sheep IgM acceptor antibody. Taking this principle one step further, in the case of an intact kappabody, the present invention contemplates the expression of a molecule having one lambda and one kappa chain, regardless of whether they were of the same species, or a molecule having two kappa or two lambda chains of different species. To insure proper disulfide bridging, heavy chain acceptor antibodies of an intact kappabody are preferably of the same species, class and subclass.

The five illustrative generalized preferred embodiments have several common, more preferred embodiments. For instance, it is preferred that the donor and acceptor antibodies for these five constructs have donor and acceptor antibodies that are different and that are chosen from murine, rabbit, or primate monoclonal or antibodies. Furthermore, it is preferred that all of the CDRs in the various CSV_L and CDR-grafted V_L domains, as the case may be, are identical in amino acid sequence to the corresponding CDRs of donor antibody CDRs; that all of the framework regions are derived from the same amino acid sequence as, (i.e., being at least about 75% and preferably at least 85% homologous to) the corresponding framework regions of the acceptor antibody(ies); and that any constant domains, whether light chain or heavy chain, as the case may be, are identical in amino acid sequence to the corresponding domains of the acceptor antibody(ies). In order to make these

preferred constructs less immunogenic, it is further preferred that the acceptor antibody(ies) be human, especially a human antibody that has light chains of the kappa class, and more so when the human heavy chains are of the gamma class. (It is understood that the class and subclass of the two heavy chains in an intact kappabody are preferably the same in order to obtain optimal disulfide bridging between the two chains.) With regard to the ScFv(CSV_L) fragment, when the acceptor antibody is human, it is preferred that the linking peptide be from about 12 to about 18 amino acid residues, and especially so when the CDR-grafted V_L domain is fused to the N-terminus of the polypeptide linker, and wherein the C-terminus of the polypeptide linker is fused to the N-terminus of the CSV_L domain.

Further preferred embodiments of this invention occur when murine monoclonal antibodies are used as donor antibodies, and more so when these murine antibodies have binding affinity, and thus were raised against, tumor antigens and antigens on thrombi; but especially so for tumor antigens of human but also of any other vertebrate origin. Preferred tumor antigens; (or markers as they are sometimes called), are AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (a multiple drug resistance marker), Proliferating Cell Nuclear Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes. An especially important tumor antigen is CEA, with preferred murine donor antibodies being the anti-CEA antibodies ZCE 025 (C. M. Haskell, *et al.*, Cancer Research, 43,3857 (1983), who refers to the antibody as "MAB 035") and CEM 231 (C.B. Beidler, et al., J. Immunology, 141,(11), 4053 (1988)). Regarding the ScFv(CSV_L) fragment, when the donor murine antibody is an anti-CEA antibody, it is

further preferred that the peptide linker be composed of serine and glycine residues.) With the latter two anti-CEA donor antibodies, it is preferred that the acceptor antibody be the human IM9 antibody, (Reference under IM-9 in ATCC #159) wherein the framework regions in the CSV_L and CDR-grafted light chain domains, as the case may be, are mostly the same in amino acid sequence as the corresponding IM9 framework regions. The most preferred donor antibody is ZCE025. Finally, with regard to the $ScFv(CSV_L)$ fragment, when the donor antibody is ZCE025, it is preferred that the peptide linker have the amino acid sequence -GGSGGSGGSGGSGG- (**Sequence I.D. No. 1**).

Each of the above five preferred embodiments can optionally have fused to its C- or N- terminus a metal-chelating peptide sequence. The chelating peptide sequence can be up to about twenty-five amino acid residues in length. In the case of the CSV_L and the $ScFv(CSV_L)$ fragment, only one such peptide chain is bound to either available terminus. In the case of the kappabody fragment and the heavybody, the chelating peptide can be bound to either one or the other, or both, chains, and when bound to both chains, can be bound to either the N-termini, the C-termini, the C-terminus of one chain and N-terminus of the other, or to both termini of both chains. With the intact kappabody, a chelating peptide such as that described above can be bound to any number of the four chains comprising the molecule, with any and all combinations of N-termini and C-termini bonding envisioned. For any one of the five preferred constructs, it is further preferred that metal chelating peptide consist of about ten amino acid residues or less and chelate to either nickel(+2), zinc(+2), copper(+2), or cobalt(+2) ions and be bonded to one or more, as the case may be, of the c-termini of the molecule. More preferred is the case where one (or more) of the C-termini is fused to a metal chelating peptide of the sequence HWHHHP (**Sequence I.D. No. 2**) through the peptide's N-terminal histidine residue. Regarding the preferred embodiment of any of the five preferred constructs

as described above, when the embodiment is narrowed to a murine donor antibody, it is preferred that a metal-chelating species be bonded to the C-terminus (or possibly more than one termini , as is applicable), consist of ten or less amino acid residues, and chelate with either nickel(+2), copper(+2), zinc(+2) or copper (+2) ions. Finally, in the most preferred embodiment of the above five constructs , that is, wherein the donor antibody is the murine monoclonal antibody ZCE 025, it is preferred that the optional metal chelating peptide have the sequence HWHHHP and be fused to the C-terminal (or one or more termini , as is applicable) of the molecule.

The present invention also comprises the RNA and DNA sequences coding for any molecule therein, including but not limited to the five preferred constructs and their corresponding preferred embodiments.

The present invention also comprises antigen-binding fragments of any of the above molecules that can be obtained by routine chemical and enzymatic manipulation, such as the fragments resulting from the chemical cleavage of bridging disulfide bonds, (e.g. using 2-mercaptoethanol and iodoacetate), and from enzymatic digestion with routine reagents such as pepsin and papain. For instance, it is within the scope of the present invention to have an $F(ab')_2$ fragment obtained from the digestion of an intact kappabody, and any of its preferred embodiments described above, with pepsin, or an Fab fragment obtained from the digestion of it with papain.

The CSV_L recombinant antibodies of the present invention contain one or more heavy chain CDR(s) from a donor antibody grafted into a kappa or lambda chain variable domain. The immunoglobulin chain containing the CSV_L can further contain either a kappa or lambda constant region, or one or more alpha, delta, epsilon, gamma or mu constant region, depending upon its intended use. As mentioned above, gamma constant regions are preferred for this invention, and

especially preferred are the constant regions of the gamma-1 subclass.

As used to define and delineate the scope of the present invention, the term "CSV_L recombinant antibodies" shall mean both a CSV_L fragment and a CSV_L-containing antibody or fragment thereof, including a Heavybody, an ScFv(CSV_L) fragment, an Intact Kappabody, and a Kappabody fragment.

In the various constructs of the present invention the antibody that provides the framework regions into which are grafted CDRs from another antibody is referred to as the "acceptor antibody." The antibody that provides the CDRs grafted into the acceptor antibody is referred to as the "donor antibody". In one embodiment, the amino acid sequence in the four framework regions of the acceptor antibody are substantially homologous (i.e. at least about 75% homology) to the corresponding regions of the native acceptor antibodies. In another embodiment, the protein sequences in the framework regions of the acceptor antibody are altered, for example, by means of computer modeling, to preserve certain amino acids from the donor antibody that are necessary to conserve the binding affinity of the CSV_L domains and the CDR-grafted light chain domain and the ability of the hybrid immunoglobulin chains containing the altered variable domains to associate and assemble with other such immunoglobulin chains into antibody-like constructs.

Since a single alteration in the protein sequence of a CDR can substantially decrease the binding affinity of the construct for its antigen, the grafted CDRs are preferably homologous to those of the donor antibody; however, it is intended that one or more residues of a donor CDR can optionally be changed or omitted. The donor and acceptor antibodies can be polyclonal or monoclonal and can be of any antibody class or species. Preferably, however, the acceptor light chains are derived from a human antibody, most preferably IgG, and the CDRs are derived from a donor

antibody from a non-human species selected from the group consisting of rodent, rabbit, and primate antibodies. Human donor antibodies may also be used and in one embodiment of the invention the CSV_L recombinant antibodies are made using the same antibody as both donor and acceptor, i.e., the heavy chain CDRs are grafted into a kappa light chain and associated with a native kappa light chain to make an engineered light chain dimer fragment.

A CSV_L recombinant antibody may have attached to it an effector or reporter molecule. For instance, a macrocycle or chelating peptide may be attached for chelating a heavy metal atom. Similarly, a toxin, such as ricin, can be attached to the recombinant antibodies of this invention by any of a number of covalent binding structures known in the art. Alternatively, a fusion protein comprising a CSV_L recombinant antibody joined by a peptide linkage to a chelating peptide or functional non-immunoglobulin protein, such as an enzyme or toxin molecule, can be produced using the procedures of recombinant DNA technology, for instance, the general methods of Neuberger, *et al.*, in PCT Patent Application No. PCT/GB85/00392.

The term "antigen" as used herein shall encompass large protein antigens, such as carcinoembryonic antigen, in addition to haptens, such as metal-binding haptens. The ability to bind with an antigen or hapten is determined by assays well known in the art, such as antibody capture assays (See, for example, Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1988)).

The CSV_L recombinant antibodies are made using techniques of genetic engineering that are well known in the art. (See for example European Patent Application EP 0 239 400 to Winter, *et al.*, PCT Patent Application PCT/GB91/-1108 to Adair, and U.S. Patent Nos. 5,132,405 and 5,091,513 to Huston, *et al.*) The terms "CDR grafted", "grafted with", and "grafted into", and the like, as used herein shall have the meaning well known in the art that, using the techniques of

genetic engineering, in one antibody, called the acceptor antibody, the CDRs are removed and replaced with those of another antibody, usually of another species, called the donor antibody. In the CSV_L recombinant antibodies taught herein a CDR from the donor antibody can be grafted into a CDR locus in the acceptor immunoglobulin other than the one from which it is derived in the donor immunoglobulin. That is, CDR1 in the acceptor immunoglobulin can be replaced with CDR2 or CDR3 from a donor antibody, and so forth. The CSV_L recombinant antibodies may comprise only one or two donor-derived CDRs, though preferably all three CDRs are derived from the donor antibody and are grafted into the acceptor frameworks so as to replace the native CDRs therein, i.e., donor CDR1 of the opposite chain is grafted into the locus of CDR1 in the acceptor immunoglobulin chain. As used herein the terms "CDR" and "framework region" shall have the meanings and their locations shall be determined according to the method of Wu and Kabat, J. Exp. Med. 132:211-250 (1970), unless crystallographic analysis or homology modeling dictate that they have slightly modified locations.

As used herein the phrase "derived from" and "altered" shall encompass the meaning that certain amino acids (less than or equal to 25% and preferably less than or equal to 15% of the total amino acid residues) in the acceptor framework regions of the CDR grafted constructs are switched to match the corresponding amino acids from the donor antibody as needed to facilitate the dual goals of preserving the binding affinity of the donor antibody and the expression levels of the acceptor antibody.

The CSV_L recombinant antibodies of this invention can be engineered to have the size, function and general design of an intact antibody or of any antibody fragments, such as Fv, Fab', single chain Fv, or single domain antibody (for example, an isolated heavy chain variable region), so long as each contains at least one CSV_L domain.

The CSV_L recombinant antibodies can be labeled for use in in vivo diagnosis and therapy. For instance radioactive ions having suitable properties for use in in vivo regimens can be attached to the recombinant antibodies under conditions similar to those known in the art.

In the kappabody (**Figure 3**), the two chains are joined by one or more, preferably one or two, sulfhydryl bridges at the C-terminus of the light chain constant domain. In the native kappa chain, there is one sulfhydryl bridge to the heavy chain, but additional sulfhydryl-bearing cysterne residues could be added by incorporating all or part of the hinge region of an IgG heavy chain or by fusing an appropriate metal-binding protein containing cysteine.

Kappa and lambda dimer fragments occur in nature and result from spontaneous combination of light chains within the host cell upon expression. Like these naturally occurring light chain dimer fragments, those of the invention associate naturally within the host cell and are held together by weak bonding interactions between the two chains, (i.e., hydrogen bonding and Van der Waals forces), by a spontaneously formed disulfide bridge at the C terminus of the chains, as well as by any natural forces of attraction of the heavy CDRs for the light CDRs. Unlike the naturally occurring light chain dimer fragments, however, it is believed that the CSV_L recombinant antibodies of the invention may experience dislocation of some of the sites of weak bonding interaction in the kappa chains (as compared to native kappa dimer fragments) due to strain caused by the splicing of foreign CDR's into the acceptor kappa chains. Therefore, in the kappabody fragments of the present invention certain residues in the acceptor framework regions holding the donor CDRs are preferably altered to overcome the effects upon affinity and specificity of the foreign CDR(s) and to ensure the ability of the engineered proteins to properly assemble upon translation.

These small (50 kd), humanized molecules offer several advantages over Fab antibody fragments. First, they are readily expressed from the same vector due to uniformity of the two chains, thus allowing for rapid construction and more equivalent expression of both chains. Second, since they are recombinantly expressed, the native carboxy-terminus is present on both chains; whereas fragments created by treatment of whole antibody with enzyme lack the native terminus and therefore can be more immunogenic. And third, these molecules, which have a structure distinct from Fab antibody fragments, are expressed at high levels and are highly stable.

It is known that during trafficking of immunoglobulin proteins within the eukaryotic cell, the heavy chain binds to the chaperon protein complex Bip/GRP94 located within the rough endoplasmic reticulum, and is thereby prevented from passage into the Golgi apparatus and thus is prevented from expression by the cell. A heavy chain is not secreted in eukaryotic cells unless or until it is displaced from the chaperon protein by a light chain, with which the heavy chain combines, thereby leading to secretion of intact antibody. For potentially similar reasons, a chimeric construct comprising the variable domain of a heavy chain and the constant region of a light chain (i.e., a V_HC_K fragment) will not be secreted by itself in mammalian host cells.

However, the instant invention discloses that a genetically engineered gene encoding a CSV_L fragment, when operably linked to the required transcriptional and translational sequences functional in eukaryotic host cells suitable for expression of immunoglobulin genes, will be transcribed, translated and secreted. The secretable CSV_L can be incorporated into constructs that also contain a light chain constant region and will convey upon the resulting the similar ability to be secreted in eukaryotic cells.

Indeed, just such a single chain fragment has been mentioned above as a preferred embodiment of this invention. A species of this "heavybody" fragment is depicted in

Figure 2. As with the kappabody fragment, two different acceptor antibodies can be used, one for the light chain constant domain and another for the FRs, although it is preferred that the same acceptor antibody be used for both areas. As illustrated below, the heavybody fragment is secreted in mammalian cells as a homodimer (an assembly of two identical chains) in the absence of the expression of native light chain by the host cell. However, in the presence of light chain, a light-heavy heterodimer (an assembly of two significantly different chains) is preferentially formed. The binding affinity of the heavybody homodimer can readily be assayed, using methods known in the art, such as a competition ELISA.

Unlike isolated native light chains of antibodies or native light chain homodimers, which do not possess binding affinity by themselves, the instant heavybodies (i.e., the single chain monomer) retain the ability to bind antigen. If it is desired to assay the binding affinity of an isolated heavybody, the sulfhydryl bridge(s) that join the chains of the heavybody homodimer can be reduced by treatment with enzyme under conditions mild enough to preserve the binding affinity of the isolated heavybody monomer using techniques well known in the art, or as is illustrated in the Examples. The heavybody is a very small (25 kd) humanized molecule of different structure from a native kappa or lambda chain. And, unlike a chimeric heavy chain, the heavybody molecule is secreted from mammalian cells with high levels of expression.

As one skilled in the art will appreciate, the present invention enables production of recombinant antibodies of smaller size. For instance, fragments analogous to Fv fragments can be made from the variable domains of two acceptor light chains by grafting at least one light chain CDR into one copy of the light chain variable region and at least one heavy chain CDR into another copy of the light chain variable domain of the donor antibody. Like Fv fragments, these smaller constructs lack the natural

sulfhydryl bridge that connects naturally occurring kappa dimers (and would leave out the light chain constant domain.)

It is also possible to rapidly engineer and secrete in mammalian host cells at high expression levels a single domain construct comprising an acceptor light chain variable region with one or more donor antibody heavy chain CDRs grafted between the framework regions referred to herein as a CSV_L fragment (as discussed above), as illustrated in **Figure 1** and described further in the Examples. This (embodiement of the present invention) further evidences that the DNA sequences effective for expression of the heavybody fragments in mammalian cells are contained in the framework regions of the light chain variable domain.

Not all antibodies or fragments with useful affinities for their antigen have heavy chain variable domains with sufficient affinity to bind with the antigen. However, by proper screening of the genome of a lymphoid cell, a heavy chain variable domain having CDRs with sufficient antigen affinity to bind as a single domain antibody can be found using techniques well known in the art. For instance, Ward, *et al.*, in "Binding Activities of a Repertoire of Single Immunoglobulin Variable Domains Secreted from *Escherichia coli*," Nature 341:544-546, (1989) disclose a method for screening to obtain an antibody heavy chain variable region (V_H single domain antibody) with sufficient affinity for its target epitope to bind thereto in the single domain format.

Alternatively, a phage expression library can be prepared from V_H DNA fragments using methods well known in the art. (See for instance, Garrard, L.J., et al., PCT Patent Application PCT/US91/09133, assigned to Genentech. Proteins expressed on the phage head can be screened using an affinity column having bound antigen or a polypeptide probe constructed from the peptide sequence of the desired target epitope or antigen. Single domain V_H antibodies that bind with the antigen can be selected and ranked to obtain those with the highest affinity for the antigen. These single

domain V_H antibodies, however, cannot be secreted in mammalian cells. By adopting these and other techniques, such as molecular modeling, known in the art and/or disclosed herein, heavy chain variable domains showing antigen-binding affinity can be obtained and used as the donor antibody to make a single domain CSV_L-containing recombinant antibody fragment according to the present invention, i.e., having one or more CDRs from a high affinity donor heavy chain variable domain grafted into the framework regions of a acceptor light chain variable domain, and preferably wherein the acceptor antibody has kappa light chains and is of human origin.

As illustrated below in the Examples, a preferred embodiment of the single domain fragment of this invention, namely the CSV_L fragments, can be expressed in mammalian cells. In contrast, a conventional single domain antibody, (i.e., one consisting of a V_H domain) cannot.

With a molecular weight of 12.5 kd, approximately one sixteenth that of intact antibody, the CSV_L fragments of the invention bind to target antigen with the specificity of the donor antibody, and with the potentially greater binding ability than the variable domain of a light chain alone. Yet these extremely small peptides will clear from the circulation more rapidly with decreased normal tissue retention and decreased immunogenicity, and penetrate tumor more extensively than any other size of antibody fragment. Even when the framework sequences of the CSV_L fragment have been altered in accordance with this invention to facilitate folding of the molecule into a three-dimensional geometry that provides the specificity and a sufficient affinity for use in <u>in vivo</u> imaging and therapeutic applications, the CSV_L fragment proteins are generally approximately thirty to thirty five percent human when three non-human CDRs have been grafted into them. Therefore, these very small recombinant fragments, which can be rapidly engineered to improve affinity or specificity due to their small, single chain format, are particularly useful for <u>in vivo</u> applications that require rapid clearance of the unbound binding fragment from

the blood, such as in vivo radiotherapy using strong beta-emitting particles attached to the binding fragment.

As mentioned above, insertion of donor CDRs into the acceptor framework regions can displace the CDRs out of their preferred spatial alignment. Association sites between the heavy and light chains can also be disrupted by introduction of the foreign CDRs so that the expression level of the CDR grafted construct is impaired relative to that of the intact acceptor antibody. In the CSV_L recombinant antibodies of the present invention, an additional problem is encountered. Grafting of heavy chain CDRs into light chain framework regions in the making of a CSV_L, can produce either different or additional dislocations of the sites in the framework regions that are necessary to support the CDRs in their preferred spatial orientations and dislocations of the association sites between the light and heavy chains that contribute to assemblage of the recombinant antibody chains during expression.

To accomplish the dual goals of (1) preserving the spatial orientation of the CDR loops as it appears in the donor antibody, and (2) maintaining to the greatest extent possible the expression levels and reduced immunogenicity of the acceptor antibody, any of a number of available methods based on computer-assisted molecular modeling procedures can be used or modified for effectively identifying and replacing amino acids in the acceptor framework regions to create CSV_L recombinant antibody of this invention.

For instance, Adair, J., et al., PCT Patent Application PCT/GB90/02017, assigned to Celltech, disclose a method for introducing mutations into acceptor framework regions of CDR-grafted antibody chains of anti-CEA antibodies to match the corresponding donor residues. In the Celltech method, in addition to the Kabat-defined CDRs from the donor antibody (CDR1: positions 24-34; CDR2: positions 50-56; CDR3: positions 89-97) the structural loop residues (positions 89-97) in CDR3 and residues at one or more of positions 1, 2,

and/or 3, 46, 47, 49, 60, 70, 84, 85 and 87 are replaced by the corresponding donor residues, if they differ.

According to the Celltech method, in the heavy chain, in addition to the Kabat-defined CDRs, the amino acid residues of the acceptor variable domain are replaced at positions 23 and 24 and 71 and/or 73 with those of the donor antibody, if they differ. Additionally, in the heavy chain, the acceptor residues can be replaced by donor residues at some or all of positions 48 and/or 49, 69, 76 and/or 78, 80, 88 and/or 91 and 96. The definitions of the CDRs can also be shifted to accommodate idiosyncratic regions in any given donor antibody.

Ideally, commercially available computer programs are used with actual crystal structures of the donor and acceptor antibodies (bound to their antigens) to determine which amino acids in the CDRs (and framework regions) contain atoms that are close enough to atoms in the amino acids of the antigen to interact.

Yet another method, generally referred to as homology modeling, is useful when a crystal structure cannot be obtained for the antibody to be used in making the antibodies of this invention. Several fully automated algorithms to align crystal structures and define structurally conserved regions are known. The loop regions are modeled by two basic methods: 1) use of a data base of available structures to provide the best possible loop conformations or, 2) use of distance-geometry based mathematical model to generate further possible conformers. The best conformer chosen by either method of modeling is chosen on the basis of some type of energy function, usually an energy calculation. For instance, computer programs such as Insight II, Homology and Discover (Biosym, San Diego, CA) are employed in conjunction with a database containing the known crystal structures of proteins, such as the Brookhaven Protein Data Bank, to construct a three dimensional representation of the immunoglobulion of interest. This three dimensional representation is based upon homology

between structurally conserved regions (SCRs) of the known structures and corresponding regions in the protein whose crystal structure is unknown. A loop search algorithm is used to identify protein loops from the database with the right number of residues and correct three dimensional disposition of backbone atoms of the regions flanking the loop to splice between the structurally conserved regions. In this way the three dimensional model of the donor and acceptor antibodies is constructed by the computer so that the residues of the donor antibody frameworks that are involved in supporting the CDRs and the residues of the acceptor antibody frameworks that are involved in chain association can be conserved in the CSV_L recombinant antibody.

The preferred method of making the CSV_L recombinant antibodies of this invention, when actual crystal structures of the donor and acceptor antibodies are not known, employs molecular modeling. Molecular modeling can be used to locate the three dimensional structurally conserved regions (SCRs) common among all antibodies. Separate computer models of the donor and acceptor immunoglobulins are constructed by a technique of homology modeling based upon a database of known protein crystal structures, such as the Brookhaven Protein Data Bank of known protein crystal structures, using the computer modeling programs Insight II, Homology and Discover, Version 2.1.2. From computer models of the donor and acceptor antibodies, the amino acid residues in each structure involved in association of the immunoglobulin chains in the acceptor antibody are determined and conserved in the CDR grafted construct. In addition, the amino acid residues involved in support of the CDRs in the donor antibody are conserved in the CSV_L recombinant antibodies.

Briefly, for the purpose of modeling the light chain variable region of an antibody, at least two and preferably at least eight antibodies are selected from a protein database, such as the Brookhaven Protein Data Bank,

that provides both a linear amino acid sequence and three dimensional atomic coordinates of each antibody variable region. The sequences and structures of these antibodies are manipulated by a computer program having the ability to assign the corresponding atomic coordinates from a segment of a known structure to the atoms of any segment of an amino acid sequence having the same number of residues. One skilled in the art will know of computer programs and databases that are suitable to work in tandem in this fashion. For example, the Brookhaven Protein Data Bank can be used together with the current versions of molecular modeling programs Insight II, Homology and Discover (Biosym Technologies, Inc., San Diego, CA); as discussed in the immediately following sections.

Step One - Definition of Structurally Conserved Regions Using Known Three-Dimensional Structures of Antibodies

Using the selected three dimensional protein structures and sequences from the database, the operator uses the computer program to align the sequences of the variable regions and to superimpose the corresponding structures so that structurally conserved regions can be identified. For instance, the sequences are aligned in a linear array, with each sequence constituting one row of the array, i.e., Seq a, Seq b, Seq c, etc.

To facilitate alignment by placing the SCRs into columns using the Insight II software, certain landmark amino acids known to be universally conserved among antibodies, such as the cysteines that form the intrachain disulfide bridge, are identified in each sequence and are aligned in vertical columns. Taking the first two of the linearly aligned sequences, one, for instance Seq a, is designated to be held constant and the other, for instance Seq b, to be superimposed onto the first. (In practice, the bottom

sequence on a computer display is usually most convenient to hold constant.)

Three-dimensional alignment of the known structures is further refined by discovering additional amino acid sequences that correspond to regions in all the selected antibodies that preserve almost identically the same three-dimensional conformation, called herein the structurally conserved regions (SCRs).

Using the already superimposed structures, the first putative SCR, conveniently designated SCR1ab, is discovered by visual inspection. Preferably, successive SCRs are identified by working from the amino to the carboxy terminus of the molecules. The RMS deviation of the backbone atoms in the two segments of amino acids corresponding to SCR1ab is calculated. The exact locations of SCR1ab, and hence of the amino acids contained within the segments corresponding to the SCR1ab, are adjusted by a procedure of trial and error whereby the amino acids in the linear sequences of the array that correspond to those in the putative SCR1ab are boxed and the RMS deviation is calculated. The width of the box is maximized and the location of the box is adjusted until the RMS deviation reaches an acceptable maximum, for instance no more than about 0.75 Angstroms.

To ensure that spatial alignment of SCR1ab at the amino terminus of first and second structures is not destroyed by establishment of subsequent SCRs along the sequences (i.e., SCR2ab and SCR3ab, etc.), preferably after the process has been carried out to define SCR2ab, the two structures are superimposed again using the residues for the backbone atoms in SCR1ab as well as SCR2ab. This process is repeated for each subsequent SCR. Gaps, for example, empty space holders, can be inserted within nonconserved (nonhomologous) regions, referred to herein as NSCRs. Usually the NSCRs are found in the loops and CDRs. Gaps are inserted as needed to accomplish vertical alignment of the

SCRs, for example, where any sequence had fewer amino acids between the SCRs than did the other.

Usually by the method of this invention, from about seven to ten SCRs are established between any two light chain variable regions, with each SCR containing from about three to twenty amino acids from each of the structures, when the RMS deviation of the backbone atomic coordinates in the SCRs is no more than about 0.75 Angstroms.

Once two of the structures have been aligned in this manner, the procedure is repeated, preferably by selecting the first structure, for instance the bottom structure in the array, to be held constant (Seq a), and discovering the SCRs between that first structure and each in turn of the other structures represented in the linear array (Seq b, Seq c, etc.) to yield SCR1ac, SCR2ac, SCR3ac, etc. and then SCR1ad, SCR2ad, SCR3ad, etc. Alternatively, of course, any other method can be used whereby segments having a common spatial conformation, such as SCRs, are located within the known three dimensional structures of from six to ten antibody variable regions. For instance, one skilled in the art will appreciate that it would be possible to locate the first SCRs in the middle of the molecules and work outward therefrom in either direction, or to begin at the carboxy terminus of molecules and work progressively towards the amino terminus. The order in which the sequences (and their structures) are compared with one another can also be varied. For instance, one skilled in the art will appreciate that it would be possible not to hold a first structure constant, and instead to align any two structures and then to chose any one of those two structures to be aligned with a third, and so forth.

When all of the structures have been compared with one another by any of the alternative methods described above, for instance when each structure in the array has been in turn superimposed and aligned with the constant first structure, as is preferred herein, the next step is to identify the consensus SCRs. A consensus SCR comprises the

residues in each linear sequence that are in the intersection of all of the individual SCRs. One skilled in the art will appreciate, however, that the technique of locating the consensus SCRs can be varied so long as structurally conserved regions (SCRs) common to all of the structures in the array are located, and so long as the RMS deviation of the coordinates corresponding to the superimposed backbone atoms in all of the structures is acceptably low, for instance no more than 0.75 Angstroms.

A similar procedure is followed to locate and fix in spatial relation to one another the SCRs common to the heavy chain variable domains of antibodies, except that the sequences used in the linear array are those of the heavy chain variable domains of the antibodies in the database whose three dimensional structures are known.

Step Two - Three-dimensional Modeling of Acceptor Fv and Identification of Chain Association Residues

Now the linear sequence of the acceptor antibody chain to be modeled is displayed as an additional row in the linear array and aligned with the sequences of the eight database antibodies as described above to discover the segments of SCRs in the acceptor chain that correspond to those in each of the boxes, using as many gap-filling spaces as needed to accomplish the vertical alignment. This process is identical for light and heavy chains. The three-dimensional model of the acceptor antibody chain can now be fabricated in segments from the consensus SCRs derived above. For each SCR1 in the linear sequence of the acceptor antibody chain, the column of SCR1s in the array is inspected to find the SCR1 with greatest sequence homology to the acceptor SCR1. The computer is used to construct the model of the acceptor SCR1 by assigning to each residue in the acceptor SCR1 coordinates corresponding to those of the selected sequence from the column of corresponding database SCR1s.

At this point, any residue in the selected SCR1 that does not match the corresponding residue in the donor SCR1 is mutated to match the residue in the acceptor SCR1, while the coordinates of all the atoms in the backbone and sidechains that correspond to those in the acceptor residue are conserved. The remaining atoms are modeled under the constraints of maintaining the same bond lengths, angles and dihedrals as those in the original database residue, i.e., for the gamma and delta carbons. The process is repeated for each of the subsequent SCRs, i.e., SCR2, SCR3, etc.

Next, the length of each segment of NSCR in the acceptor chain sequence, i.e., the spanning sequence between each successive pair of boxes, is determined. Progressively from the amino terminus of the chain, NSCR segments of the acceptor chain are modeled by selecting loops from the protein database to span between the endpoints of the SCRs of the acceptor chain model constructed above. The actual number of amino acid residues in each NSCR is counted (ignoring the space-filling gaps used to accomplish vertical alignment). For each span individually, the computer is instructed to search the protein database, for instance using the Loop Search algorithm as is well known in the art, to discover from about eight to twelve candidate amino acid sequences having (1) the same number of amino acids as the actual acceptor NSCR and (2) flanking regions with the same relative atomic coordinates as the flanking SCRs in the acceptor chain model as determined above. As one skilled in the art will appreciate, depending on local structural details, either all or some subset of the residues adjacent to the loop in each SCR box can be identified as the flanking residues. The candidate sequences whose flanking regions are best fits with the relative atomic coordinates of the SCRs of the acceptor chain model, as determined by computer algorithm, are selected.

It has been discovered that in antibodies the general spatial conformation of the loops and NSCRs is

conserved. Therefore, the best candidate for NSCR1,2 in the model should have a three dimensional spatial conformation generally similar to that of the corresponding NSCR1,2 segments in the antibody database structures. For each candidate that meets this general requirement, the backbone atoms of the flanking residues of the candidate NSCR are superimposed on the backbone atoms of the corresponding flanking residues of the SCRs of the model that flank the NSCR under consideration. For example, to consider the candidates for the NSCR1,2 position in the model, the backbone atoms of the flanking residues of the candidate NSCR are superimposed on the backbone atoms of the corresponding flanking residues of the SCR1 and SCR2 sequences from the model, and the candidate having (1) the best RMS fit of the backbone atoms of its flanking residues with backbone atoms of the corresponding flanking residues from SCR1 and SCR2 and (2) a spatial orientation most like that of NSCR1s of the database antibodies displayed on the computer screen (to rule out interference with other loops) is selected. By repeating this procedure at each NSCR position, i.e. at NSCR1,2; NSCR2,3; NSCR3,4, etc., the acceptor NSCRs are selected and then placed into the acceptor model as follows.

Once the best spatial orientation for an amino acid sequence of the given loop length for each NSCR is selected, the coordinates of the backbone of the candidate segment are assigned by the computer to the corresponding NSCR in the model. Now any residue in the selected candidate sequence NSCR dissimilar to the corresponding residue in the actual sequence of the acceptor NSCR is mutated to match the acceptor sequence while the computer algorithm is used to (1) maintain the coordinates of all the atoms common between the two, and (2) model the dissimilar atoms while constraining the bond lengths, angles and dihedrals to those of the candidate residue.

Once all of the NSCRs making up the model are in turn selected from the database, fixed in space, and modeled to transform them into the coordinates of the corresponding

acceptor NSCRs, the splice regions where the SCRs join the NSCRs are preferably refined to relieve any strain in the model that results from joining the SCRs and NSCRs. This refinement can be accomplished using any suitable computer algorithm, for instance the "Repair" algorithm in Insight II, to assign the proper bond lengths, bond angles, and omega values to the residues at the splice junctions.

Now, the model as a whole is relaxed using a suitable computer algorithm to relieve any strain occasioned by the above procedures. Preferably the "Relax" algorithm of Insight II is applied in a series of sequential steps to the model as a whole. Preferably, the order of the steps is to apply the algorithm: (1) to the side chains of the NSCRs to assign proper geometries and remove any unfavorable non-bonded contacts between side chain atoms and other atoms in the molecule, (2) to all atoms of the NSCRs to remove any remaining unfavorable contacts between the NSCR and other atoms in the molecule, (3) to the mutated side chains of the SCRs to remove any unfavorable non-bonded contacts between mutated SCR side chain atoms and other atoms in the molecule, and (4) to all of the side chain atoms of the SCRs to remove the remaining unfavorable sidechain contacts.

Finally, an energy minimization procedure is performed using techniques well known in the art, for instance, using the "Discover" subprogram of Insight II, to allow the model to assume an energetically favorable conformation. In the preferred embodiment, however, the energy minimization is performed in a series of sequential steps. The entire model is first subjected to energy minimization with backbone atoms tethered to their starting coordinates with a force constant of 100 kcal/Å^2. Then an energy minimization is performed for the entire model without the backbone atoms being tethered. The result of carrying out these steps is a model of the variable domain of each of the acceptor chains.

In the method of this invention, the model of the acceptor Fv is made by the following steps: (1) identify

potential chain association residues by comparison of the sequence of the acceptor chain with the linear array of known structures and select an appropriate known structure to use in modeling chain association of the acceptor molecule, (2) make a preliminary model by superimposing the backbone atoms of the potential chain association residues of the selected known structure, (3) subject the entire molecule to energy minimization, first, with the backbone atoms being tethered to their initial coordinates and, second, without the backbone atoms being tethered, (4) identify the chain association residues in the final acceptor Fv model, excluding all residues that are part of a CDR.

In the first step, in one chain of the structures of each of the database antibodies, each residue in the variable region of that chain having an atom within 4.5 Angstroms of an atom in a residue in the other chain is identified. If the residues so identified in each database antibody are not part of a CDR and are likely to have a significant interaction with residues in the other chain, they are earmarked in the linear sequence of the antibody as chain association residues. The process is repeated for the other chain of each database antibody.

Since the residues involved in chain association are generally conserved among antibodies, it can be assumed that there will be homology between chain association residues in the V_L and V_H of the database antibodies and those in the VL and VH of the acceptor antibody or immunoglobulin. Therefore, when a residue in the acceptor sequence is found to be identical to one earmarked in the array, it is earmarked as a chain association residue in the acceptor model. On the other hand, when an amino acid is found in the acceptor that differs from the corresponding one in the database antibody in any of the positions in the database sequences earmarked as chain association residues, it is designated as potentially disruptive to chain association. Each database antibody is compared with the acceptor molecule. The database antibody with the greatest

excess of favorable residues over disruptive residues is chosen.

In the second step, superimposition is accomplished using a program such as the "Superimpose" command in Insight II.

In the third step, a program such as "Discover" in Insight II is used to carry out the energy minimization, with the back bone atoms being tethered to their initial coordinates with a force constant (usually 100 kcal/Å^2) for the initial minimization and with no tethering for the final minimization.

In the fourth step, chain association residues in the light chain are identified as all residues from the light chain that contain an atom within a specific distance of any atom of any residue in the heavy chain selected as indicating possibility of significant interaction there between (usually about 4.5Å). Similarly, chain association residues in the heavy chain are identified as all residues from the heavy chain that contain an atom that is within a specific distance of any atom of any residue in the light chain selected as indicating possibility of significant interaction there between (usually about 4.5Å) .

Step Three - The Three-dimensional Modeling of Donor Fv and Identification of CDR Associated Residues

Models of donor Fv are arrived at in a manner identical to that described above for the acceptor Fv.

CDR-associated residues are identified after minimization by determining those residues containing an atom within a specific distance of any atom of any residue found within a CDR selected as indicating the possibility of interaction there between (usually about 4.5Å). These residues are defined as CDR-associated residues and are treated in a step in the humanization process described in Step 4 below.

Step Four - Three-dimensional Modeling of CDR-grafted Fv.

The CDR-associated residues determined above are now identified in the primary amino acid sequence of the donor molecule, and the primary sequence for the altered light and heavy chain CDR grafted molecules are pieced together in segments.

First, the primary amino acid sequences of the donor and acceptor molecules are aligned with reference to the sequences of the known database structures. Second, on the donor linear array: (1) the CDR-associated residues determined above are identified, (2) for SCRs or NSCRs that do not contain a CDR residue or a CDR-associated residue, the sequence of the entire segment is replaced with the sequence from the corresponding segment of the acceptor molecule, (3) for SCRs that contain one or more CDR residues or CDR-associated residues, all residues that are neither CDR nor CDR-associated in the segment are replaced with those of the acceptor molecule, but the CDR residues and CDR-associated residues are conserved as the donor residues, (4) in NSCRs that contain one or more CDR residues or CDR-associated residues, if the total number of residues in the NSCR differs between the donor and acceptor, the entire NSCR is conserved as the donor sequence, (5) in NSCRs that contain one or more CDR residues or CDR-associated residues, if the total number of residues in the NSCR is the same between the donor and acceptor, those residues that are neither CDR nor CDR-associated are replaced with those of the acceptor molecule, while the CDR residues and CDR-associated residues are conserved as the donor residues. Thus, in all cases, CDR residues and CDR-associated residues in SCRs or NSCRs are conserved as the donor residues.

Third, the donor and acceptor models are superimposed. Once the two models are brought up on the computer screen, SCRs are determined. In this step SCRs are derived in a way distinct from that used in construction of

the acceptor and donor models. In the latter case the SCRs were assigned to the donor and acceptor based on the consensus SCRs determined from the known structures. In this step, SCRs are determined anew from the two models alone in a manner analogous to that used to determine the SCRs between each of the known structures, as described in **Step 1** above (wherein the acceptor was designated to be held constant and the donor was superimposed upon it).

Using the modeled three dimensional structures and sequences for the acceptor and donor Fvs, the operator uses the computer program to align the sequences for the Fvs and to superimpose the corresponding structures so that SCRs can be identified. For instance, the sequences are aligned in a linear array with each sequence constituting one row of the array, i.e. seqA (for acceptor) and seqD (for donor). To facilitate alignment using the Insight II software, certain landmark amino acids known to be conserved among antibodies, such as the cysteines that form the intrachain disulfide bridge (i.e., the light chain cysteines at L23), are identified in each sequence and are aligned in vertical columns, as described in **Step 1** above.

Three dimensional alignment of the two structures is further refined by identifying SCRs and superimposing them. Using the already superimposed structures, the putative SCR1AD is discovered by visual inspection. Preferably, successive SCRs are identified by working from amino to carboxy terminus of the molecules. The RMS deviation of the backbone atoms in the corresponding segments of amino acids in the two structures is calculated. The exact location of SCR1AD, and hence of the amino acids contained within the segments corresponding to SCR1AD, are adjusted by a procedure of trial and error whereby the amino acids in the linear sequences of the array that correspond to those in the putative SCR1AD are boxed and the RMS deviation is calculated. The width of the box is maximized and the location of the box is adjusted until the RMS deviation

reaches an acceptable maximum, for instance no more than about 0.75Å.

To ensure that spatial alignment of SCR1AD at the amino terminus of the two structures is not destroyed by establishment of subsequent SCRs along the sequences (i.e., SCR2AD, SCR3AD, etc.), after the process has been carried out to define SCR2AD, the two structures are superimposed again using the residues for the back bone atoms in SCR1AD as well as SCR2AD. This process is repeated for each subsequent SCR.

Gaps, for example empty space holders, can be inserted within NSCRs as needed to accomplish vertical alignment of the SCRs For example, where any sequence has fewer amino acids between the SCRs than does the other, gaps can be used to make the two of equal length.

Each segment in the altered CDR grafted chain is assigned spatial coordinates that correspond to those of the donor or acceptor residue to which it corresponds. Preferably this is done working from the amino to the carboxy terminus of the chain.

Now the light and heavy chain minimized models constructed above are displayed on the computer screen together as an Fv. An energy minimization is performed to allow this Fv model to assume an energetically favorable conformation using the steps described above.

As a final check, the model is examined to determine whether any new CDR-associated residues appear in the altered, CDR-grafted model using the techniques described above. If any new CDR-associated residue is seen in the altered CDR-grafted (and humanized) model, the amino acid at that position is replaced by the one found in the donor molecule. After the CDR-associated residues are modified as necessary, the model is analyzed to determine whether all the chain association sites identified in the acceptor model have been conserved in the altered CDR-grafted model. If differences are observed, they should be noted as possible future sites for mutagenesis if a significant decrease in

secretion of the altered CDR-grafted protein is observed as compared to that of the acceptor molecule.

Step Five - **The Three-dimensional Modeling of CSV_L/CDR-grafted-V_L Fv**

The acceptor light and donor heavy chain primary amino acid sequences had already been aligned with reference to different sequences. Therefore, it was necessary to bridge these alignments through realignment using a common sequence. In addition, the acceptor heavy chain provided information on chain association residues. Donor heavy chain sequence was added to a linear array containing light chain donor and light and heavy chain acceptor sequences and aligned. Once aligned in this manner, SCRs were defined there between as described in **Step One**, the Kabat defined CDRs and CDR-associated residues determined in **Step Three** were identified on the donor heavy chain linear array. For SCR or NSCR regions which do not contain a CDR or CDR-associated residue, the entire region was replaced with the acceptor light chain sequence (and structure, i.e., coordinates). For SCRs which contain one or more CDR or CDR-associated regions, the non-CDR-associated residues were replaced with acceptor sequence (and structure, i.e., coordinates), but donor heavy chain sequence (and structure, i.e., coordinates) was conserved for the CDR-associated residues. For NSCRs that contain one or more CDR or CDR-associated residues, the donor heavy chain sequence (and structure, i.e., coordinates) was conserved for the entire region. In this way the primary sequence for the heavy chain CDR-grafted molecule was determined, and a composite structure was developed.

Now, the resultant model was modified to assure that chain association residues, derived from the acceptor model were conserved. In all non-CDR or non-CDR-associated regions, when the amino acid in the position occupied by the chain association residue was different than the

corresponding acceptor residue, it was replaced with the corresponding acceptor heavy chain chain-association residue. In this example no chain-association residues were found to lie in the CDR or CDR-associated regions. In the unlikely event that this should occur, the residue should be noted, but no change should be made.

Alternatively, humanized light chain can be used as acceptor and humanized heavy chain can be used as donor. In this case, chain association residues used for the preliminary Fv model are those identified for humanized FV.

Now that coordinates had been assigned for both light and heavy/light hybrid chains, these were displayed on the screen together. An energy minimization was performed using the "Discover" subprogram to allow the model to assume an energetically favorable configuration. First the entire model was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a force constant of 100 Kcal/Å^2. Then the energy minimization algorithm was applied to the entire model without the backbone atoms being tethered.

CDR-Associated residues were determined for the modeled humanized light chain dimer as for the original donor Fv **(Step Three)**. Again, this was done by first identifying all residues on the light or heavy/light hybrid chain that are within 4.5 Å of any light chain CDR residue, and that also have a significant likelihood of interaction, based on orientation of the residue, charge, hydrophobicity, etc. Next, all residues on the light or heavy/light hybrid chain that were within 4.5 Å of any heavy/light hybrid chain CDR residue were identified. Again, the set was limited to those with a high likelihood of significant interaction with the CDR residue of interest. In this way, the entire set of light and heavy chain CDR-associated residues was determined.

The set of CDR-associated residues determined for the humanized light chain dimer was compared to that determined for the donor Fv. In any case where an additional CDR-associated residue is seen for the humanized, the amino

acid at that position was replaced by the amino acid found in the murine donor.

After the CDR-associated residues were modified as necessary, the model was analyzed to determine if the chain association residues identified for acceptor were conserved. In this example, they were conserved. If, however, differences are observed, these are noted, but no changes are made at this time. If, in addition, there is a significant decrease in expression observed for the humanized molecule, these are potential sites for modification.

As can be seen from the results presented in the Examples below, these modeling methods yield high affinity CSV_L recombinant antibodies from an initial design without a requirement for iteration.

This method of modeling CDR switched antibodies using structurally conserved regions can readily be modified by one skilled in the art to produce the CSV_L recombinant antibodies of this invention, such as heavybodies or CSV_L fragments.

The acceptor amino acids identified as candidates for switching to donor amino acids by molecular modeling can be switched by oligonucleotide directed or site-directed mutagenesis of the DNA sequences encoding the CDR grafted heavy and light variable regions, for instance, as taught by T. Kunkel, Proc. Natl. Acad. Sci. USA, 82:488-492 (1985) or by codon-based mutagenesis whereby an amino acid alteration is obtained for each in vitro substitution of a three nucleotide codon (Huse, *et al.*, Science, 246:1275 (1989)). Preferably however, the DNA of the entire variable region of the heavy and light chains is prepared by oligonucleotide synthesis as described hereafter.

Once the DNA encoding a CSV_L recombinant antibody has been prepared, it is then incorporated into a vector and operably linked to nucleic acid sequences encoding transcriptional and translational regulatory sequences. Any suitable expression vector may be used in this invention and exemplary vectors are provided in the Examples below. Those

with skill in the art will appreciate that the choice of vector is limited to those vectors capable of directing expression of the nucleic acid sequence encoding protein and to those vectors that can incorporate and support the function of the regulatory regions used. Further, the choice of vector is limited by the cell type selected. Not all vectors and not all regulatory elements necessary for recombinant protein expression function in all cell types. As a general rule eukaryotic expression vectors are suitable for protein expression in eukaryotes and prokaryotic expression vectors are suitable for prokaryotes. Both types of vectors are commercially available and those with skill in the art of molecular biology will be able to select appropriate vectors suitable for recombinant protein expression within a given cell type.

Methods for incorporating a particular region of nucleic acid into a nucleic acid vector are well known in the art of molecular biology (See Sambrook, *et al.*, Molecular Cloning A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989). For example, short regions of nucleic acid (less than 400 bp) can be prepared by generating sense and antisense oligonucleotides complementary to the desired gene sequence that overlap. These oligonucleotides hybridize to one another, and can be amplified in a PCR reaction, ligated and incorporated into an appropriate expression vector (see generally H.A. Erlich, PCR Technology: Principles and Applications for DNA Amplification, W.H. Freeman and Co., New York, 1992).

In general, the recombinant antibodies of this invention can be prepared by recombinant methods known in the art (see generally, Sambrook, *et al.*, supra) from the amino acid and DNA sequences of the donor and acceptor antibodies. For instance, if a monoclonal antibody is used as the donor antibody, hybridoma or polydoma technology using conventional procedures for immunization of mammals with an immunogenic antigen preparation, fusion of immune lymph or spleen cells

with an immortal myeloma cell line, and isolation of specific hybridoma clones can be employed to obtain the monoclonal antibody.

Alternatively, the genes encoding the donor and acceptor antibodies can be obtained by methods known in the art, for instance by chemical synthesis, as described above, if the sequences of the genes are known. If the sequences are not known, or if the genes have not previously been isolated, they may be cloned from a cDNA library (made from RNA obtained from a suitable tissue or batch of cells in which the desired gene is expressed, such as a hybridoma or polydoma) or from a suitable genomic DNA library. The mRNA is extracted and cDNA for the coding regions is derived using the enzyme reverse transcriptase and methods well known in the art. The gene is then identified using an appropriate molecular probe. For cDNA libraries, suitable probes include monoclonal or polyclonal antibodies (provided that the cDNA library is an expression library), oligonucleotides, and cDNAS or fragments thereof. The probes that may be used to isolate the gene of interest from genomic DNA libraries include cDNAS or fragments thereof that encode the same or a similar gene, homologous genomic DNAs or DNA fragments, and oligonucleotides. Screening the cDNA or genomic library with the selected probe is conducted using standard procedures as described in chapters 10-12 of Sambrook, *et al.*, supra.

From the sequence of the cDNA or that of the genomic DNA, the corresponding amino acid sequences to be used in molecular modeling are deduced, usually by a computer software program, such as is commercially available from DNAStar (Madison, WI). Once the amino acid sequences of the donor and acceptor antibodies are known, their CDRs are identified using the procedure of Kabat and Wu, supra. For modeling and construction of a CSV_L domain, the amino acids corresponding to at least one and preferably all three CDRs of the acceptor VL are replaced with CDRs of the donor VH. Additional donor residues identified by molecular modeling as

useful for retaining binding affinity and/or chain association are determined as described above.

When a nucleotide sequence capable of encoding the CDR grafted CSV_L domain has been determined from the protein sequence, it is fabricated and ligated into a suitable replicable expression vector, optionally along with the desired constant region genes from the acceptor antibody. A similar procedure is then followed to construct the vector containing the genes encoding the associated CDR grafted light chain or heavy chain if applicable using methods well known in the art.

It is preferred that the DNA encoding the entire CDR-grafted variable regions, including the CSV_L domain(s) be inserted into an appropriate sequencing vector (e.g. a TA vector) and sequenced employing, for instance, the SequenaseII kit (United States Biochemical, Cleveland, OH) used with a Genesis® 2000 automated DNA sequencer (Dupont, Wilmington, DE) according to the manufacturer's instructions. The spliced and sequenced exon is then excised from the sequencing vector and ligated into a vector that may optionally contain one or more exons encoding constant regions for the CDR-grafted chain. If it is desired to produce an recombinant antibody having a light and a heavy chain, the DNA encoding the light chain can be spliced into one vector and the DNA encoding the heavy chain can be spliced into another vector. Alternatively, the DNA encoding both chains can be spliced into the same vector.

To obtain the recombinant antibodies of the invention, the DNA encoding one or more immunoglobulin chains prepared as described above is ligated into a replicable expression vector so as to be operably linked to transcription regulatory element(s); suitable host cells are transfected with the vectors; and the transformed host cells are cultured under conditions favorable for forming the desired recombinant antibodies.

Various types of vectors may be used such as plasmids and viruses, including animal viruses and

bacteriophages. In the embodiment, a vector is employed which is capable of integrating the desired gene sequences into the host cell chromosome. The cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more marker genes which allow for selection (i.e., growth of the cells in the presence of a toxic drug) of host cells which contain the expression vector. The introduced marker gene sequence will be incorporated into the plasmid or viral vector containing the gene(s) encoding the construct containing a CSV_L domain. Factors of importance in selecting a plasmid or viral expression vector include the ease with which recipient cells that contain the vector may be recognized and selected; the number of copies of the vector which can be introduced or desired in a particular host; and whether it is desirable to "shuttle" the vector between host cells of different species.

Eukaryotic expression vectors for yeast or mammalian cells, as well as prokaryotic expression vectors, may be used to express the recombinant antibodies of this invention.

Although, either eukaryotes or prokaryotes can be used as host cells for this invention, the modeling methods used are exceptionally appropriate for eukaryotic cells, and more specifically for mammalian B lymphocytes. Alternatively, however expression can be obtained in a multitude of species, using suitable vectors and hosts. Suitable prokaryotic host cells include *E. coli* strain JM 101, *E. coli* K12 strain 294 (ATCC No. 31,336), *E. coli* strain W3110 (ATCC No. 27,325), *E. coli* X1776 (ATCC No. 31,537), *E. coli* XL-1-Blue (Stratagene), and *E. coli* B; however, many other strains of *E. coli*, such as HB101, NM522, NM538, MN539, and many other species and genera of prokaryotes may be used as well. In addition to the *E. coli* strains listed above, bacilli such as Bacillus subtilis, other enterobacteriaceae such as Salmonella typhimurium or Serratia marcesans, and various Pseudomonas species may all be used as hosts. As is well known to one skilled in the art, it is necessary to

remove any introns from eukaryotic genes which are to be expressed in prokaryotic hosts.

When the vector is designated for expression in baculovirus, suitable promoters and enhancer sequences include, but are not limited to AcMNPV polyhedrin, AcMNPV ETL and AcMNPV p10 sequences. One particularly suitable polyadenylation signal is the polyhedrin AcMNPV. Ig Kappa, Ig Heavy and AcMNPV are examples of suitable signal sequences. These vectors are useful in the following insect cell lines, among others: SF9, SF21 and High 5.

Alternatively, the polypeptides can be expressed in yeast strains such as PS23-6A, W301-18A, LL20, D234-3, INVSC1, INVSC2, YJJ337. Promoter and enhancer sequences such as gal 1 and PEFT-1 are useful. Vra-4 also provides a suitable enhancer sequence. Sequences useful as functional "origins or replication" include arsl and 2μ circular plasmid.

Following procedures outlined above, mammalian cell lines such as myeloma (P3-653) or hybridoma (SP2/0), Chinese Hamster Ovary (CHO), Green monkey kidney (COS1) and murine fibroblasts (L492) are suitable host cells for expression. These "mammalian" vectors can include a promoter, an enhancer, a polyadenylation signal, signal sequences and genes encoding selectable markers including, but not limited to, geneticin (neomycin resistance), mycophenolic acid (xanthine guanine phosphoribosyl transferase) or histidinol (histidinol dehydrogenase).

Suitable promoters for use in mammalian host cells include, but are not limited to, Ig Kappa, Ig heavy, Cytomegalovirus (CMV) immediate early, Rous Sarcoma Virus (RSV), Simian virus 40 (SV40) early, mouse mammary tumor (MMTV) virus and metallothionein. Suitable enhancers include, but are not limited to Ig Kappa, Ig Heavy, CMV early and SV40. Suitable polyadenylation sequences include Ig Kappa, Ig Gamma or SV40 large T antigen. Suitable signal sequences include, but are not limited to, Ig Kappa, Ig Heavy and human growth hormone (HGH).

For expression in mammalian cells the vectors containing the DNA encoding the heavy and light chain genes of the antibody construct can be placed into separate bacterial amplification vectors, such as *E. coli* DH 10 B Electromax (BRL, Gaithersburg, MD.), cultured, and screened for antibiotic resistance to amplify the plasmid. Generally, the DNA of the selected clones is verified by restriction digestion and DNA sequencing. Double stranded dideoxy sequencing is performed, for example on a DuPont Genesis® 2000 instrument, using the DuPont Genesis® 2000 sequencing kit according to the manufacturer's instructions. Post gel processing can be done with the Base Caller 5.0 program (DuPont, Boston, MA). One skilled in the art can readily provide alternative methods of performing these steps in the cloning process.

Particularly useful vectors for expression of the CSV_L recombinant antibodies of this invention in mammalian cells are pGIM9kappa and pNIM9k/hCEM-gamma deposited with the ATCC under the requirements of the Budapest Treaty under Accession Nos. 75512 and 75511, respectively. These vectors comprise human immunoglobulin regulatory elements and contain cassette sites for insertion of DNA encoding CDR grafted light and heavy chain sequences. These vectors, which are especially designed for expressing CDR grafted antibodies and fragments wherein the acceptor antibody is human, are preferably transfected into host cells of the B-cell lineage for production of optimal levels of immunoglobulin. Use of these vectors is exemplified in the examples below. The principal advantage of expressing the CSV_L domain in the above described vectors in host cells of the B-cell lineage, is that this allows for maximal conservation of assembly and secretory components to assure reproducible high level expression and secretion of the molecules of interest.

After selection of the transformed cells, these cells are grown in culture media and screened for expression of the appropriate antibody construct using techniques well known in the art for enzyme or radio assay, or by the methods

exemplified in **Example 15** below. Expression of the sequence results in the production of the fusion protein of the present invention.

A chelator may also be bound to the CSV_L recombinant antibody through a short or long chain linker moiety, through one or more functional groups on the antibody, e.g., amine, carboxyl, phenyl, thiol or hydroxyl groups. See for example Schlom, "Monoclonal Antibody-based Therapy of a Human Tumor Xenograft with a ^{177}Lutetium-labeled Immunoconjugate," Cancer Research, 31:2889-2896 (1991); U.S. Patent 4,994,560 to Kraper, *et al.*; and Sigel, *et al.*, "Coordinating Properties of the Amide Bond. Stability and Structure of Metal Ion Complexes of Peptides and Related Ligands," Chemical Review, 82:385-426 (1982). Various conventional linkers can be used, e.g., diisocyanates, diisothiocyanates, carbodiimides, bis-hydroxyxuccinimide esters, maleimide-hydroxysuccinimide esters, glutaraldehyde and the like, for instance, a selective sequential linker such as the anhydride-isothiocyanate linker disclosed in U.S. Patent 4,680,338.

This invention also contemplates fusing at least one of the genes encoding the CSV_L recombinant antibodies to a second gene encoding a chelating peptide for binding a radiometal ion, a toxin, or an enzyme such that a fusion protein is generated during transcription and translation. Fusion of two genes may be accomplished by inserting the gene encoding the chelating peptide into a particular site on a plasmid that contains an antibody gene, preferably a constant region gene, or by inserting an antibody gene into a particular site on a plasmid that contains a gene encoding the chelating peptide.

The plasmid is cut at the precise location that the gene is to be inserted using a restriction endonuclease site (preferably a unique site). The plasmid is digested, phosphatased, and purified as described above. The gene encoding the second protein or protein segment is then inserted into this linearized plasmid by ligating the two

DNA's together such that the reading frames of the gene already in the plasmid and of the gene to be inserted are preserved. If the two pieces of DNA to be ligated have blunt ends or sticky ends, ligation can be direct using a ligase such as bacteriophage T4 DNA ligase and incubating the mixture at 16°C for 1-4 hours or overnight in the presence of ATP and ligase buffer as described in Section 1.68 of Sambrook, *et al.*, supra. If the ends are not compatible, they must first be made blunt by using the Klenow fragment of DNA polymerase I or bacteriophage T4 DNA polymerase, both of which require the four deoxyribonucleotide triphosphates to fill in overhanging single-stranded ends of the digested DNA.

When constructing a replicable expression vector containing the DNA, encoding one or more of the chains of the instant CSV_L recombinant antibodies, all subunits can be regulated by the same promoter, typically located 5' to the DNA encoding the subunits, or each can be regulated by a separate promoter suitably oriented in the vector so that each promoter is operably linked to the DNA it is intended to regulate. When the CSV_L DNA is composed of subunits, for example, the DNA for the heavy and light chains of an intact kappabody, generally one of the subunits is fused or operably linked to the gene for the chelating peptide, if one is included. This fused gene will contain a functional signal sequence. A separate gene encodes the other subunit or subunits, and each subunit generally has its own signal sequence. Alternatively, to increase the specific activity of the gene fusion product, more than one gene for the chelating peptide can be fused to a subunit. For example, the gene for the chelating peptide can be fused to the genes encoding both the heavy and light chains of any antibody or antibody fragment, such as an intact kappabody or a heavybody or Fab-like fragment. A single promoter can regulate the expression of both subunits, or each subunit can be independently regulated by a different promoter. Thus, generally the complementary chain needed to provide the binding domain of the protein ligand may be provided by

expressing the complementary chain as a single polypeptide in the host cell or such a single polypeptide can be added separately. For example, to produce a fusion protein composed of a chelating peptide and an kappabody fragment, a gene encoding a light chain (or portion thereof) is functionally linked to the chelating peptide gene and this hybrid gene is expressed in a host cell. To allow formation of the binding domain or double chain fragment (e.g., kappabody fragment or ScFv (CSV_L), the same host cell can be engineered to express the other chain and excrete the assembled fragment having the chelating peptide attached to the corresponding light chain. In another embodiment, the chelating peptide can be attached to the light chain and expressed alone as a fusion protein, (such as with a CSV_L or heavybody fragment) or both chains can be attached to chelating peptides as fusion proteins and the dimer construct can be expressed from a single host cell.

The molecules of this invention can be used in all in vitro diagnostic, in vivo diagnostic, and therapeutic applications for which antibodies have been used or their use proposed. These include naked antibody therapy (both those requiring effector function and those only requiring binding function), radioimmunotherapy, in vivo radioimmunodiagnostics, in vitro radioimmunometric assays, ELISA assays, quantitative ELISA assays, and immunohistochemical applications.

The scintigraphic imaging method of the invention is practiced by injecting a warm-blooded animal preferably a mammal, and more preferably a human, parenterally with an effective amount for scintigraphic imaging of the radiolabeled monospecific or multispecific antibody agent conjugate. By parenterally is meant, e.g. intravenously, intraarterially, intrathecally, interstitially or intracavitarily. For imaging cardiovascular lesions, intraveneous or intraarterial administration is preferred.

Labeling with either Iodine-131 or Iodine-123 is readily effected using an oxidative procedure wherein a

mixture of radioactive potassium or sodium iodide and the antibody is treated with chloramine-T, e.g., as reported by Greenwood, *et al.*, Biochem. J., 89:114 (1963) and modified by McConahey, *et al.*, Int. Arch. Allergy Appl. Immunol., 29:185 (1969). This results in direct substitution of iodine atoms for hydrogen atoms on the antibody molecule. Alternatively, lactoperoxidase iodination may be used, as described by Feteanu, "Labeled Antibodies in Biology and Medicine," page 302 (McGraw-Hill Int. Bk. Co., New York, 1978), and references cited therein.

Feteanu also discloses a wide range of more advanced labeling techniques, supra, pages 214-309. Introduction of various metal radio-isotopes may be accomplished according to the procedures of Wagner, *et al.*, J. Nucl. Med., 20:428 (1979); Sundberg, *et al.*, J. Med. Chem., 17:1304 (1974); and Saha *et al.*, J. Nucl. Med., 6:542 (1976), for instance.

As used in the methods of the present invention, the compounds taught herein can be administered to the subject animal such as a laboratory animal, a mammal or more preferably a human, by any means known to those skilled in the art, including parenteral injection or topical application. Injection can be done intravascularly, intraperitoneally, subcutaneously or intramuscularly. For parenteral administration, the compounds can be administered in admixture with a suitable pharmaceutically acceptable carrier. As used herein the term "pharmaceutically acceptable carrier" encompasses any of the standard pharmaceutical carriers, such as a phosphate buffered saline solution, water, and emulsions, such as an oil/water or water/oil emulsion, and various types of wetting agents.

This invention also provides pharmaceutical compositions containing any of the CSV_L recombinant antibodies fused to the metal chelating peptides described herein linked to protein ligands, with or without the radioion having been incorporated into the chelating peptide.

Therapeutic formulations of the compositions of this invention are prepared for storage by mixing the metal chelate-protein complex with optional physiologically acceptable buffers and carriers, excipients, or stabilizers, (Remington's Pharmaceutical Sciences, 16th edition, Osol, A., Ed., (1980)), in the form of lyophilized cake or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidents including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides, proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; and the like. These pharmaceutical compositions are used for in vivo diagnostic or therapeutic purposes.

The recombinant antibodies of this invention are present in the pharmaceutical composition in an effective amount. Methods of determining effective amounts are known to those of skill in the art and depend upon a variety of factors, including the type of disorder, age, weight, sex and medical condition of the animal or human patient, the severity of the condition, the route of administration, and the type of diagnostic or therapeutic treatment desired. A skilled veterinarian or physician can readily determine and prescribe the effective amount of the compound or pharmaceutical composition required to diagnose or treat the animal or patient, respectively. Therefore, the dose of the diagnostic compound would be selected to accommodate this requirement. For diagnostic applications a typical radiodose is between 20 and 30mCi. For instance if the CSV_L recombinant antibody is an Fab' kappabody fragment the dosage is generally in the range between about 1 and 3.0mCi per nmol of fragment. As one skilled in the art will appreciate, the amount and type of CSV_L recombinant antibodies used will affect the pharmacokinetics of the compound and one skilled in the art would take these considerations into account in

selecting the proper compound and dosage in use. Conventionally, for therapeutic application, one skilled in the art would employ relatively low doses initially and subsequently increase the dose until a maximum safe response is obtained. The specific activity of the compound will determine the amount of the compound administered and hence, the dosage of the compound containing the radioion administered.

For human therapeutic regimens the typical dosage of the radioion per injection is in the range from about 10 to 30mCi per injection and the typical corresponding antibody dose is in the range from about 2 to 10mg. Although in certain instances a single therapeutic dose can be effective, more typically the patient to be treated will be administered a series of gradually increasing doses at intervals spaced appropriately to accommodate the needs of the patient. For instance, when CSV_L recombinant antibody is a kappabody fragment, is tumor-specific, and is fused to a chelating peptide incorporating Yttrium-90 as the therapeutic radioion, a typical dosage regimen would consist of repeated administration of the therapeutic compound over appropriately spaced intervals, for instance of two weeks duration, beginning with a dosage of 10mCi/2mg of antibody and increasing to a dosage of about 30mCi/10mg of antibody. If the CSV_L recombinant antibody is incorporated into a compound containing a separate chelating peptide, the weight of the chelating peptide is negligible in comparison to the weight of the antibody so that its weight can be ignored in calculating the proper ratio of radionuclide to delivery agent (i.e., chelating peptide plus antibody).

Alternatively, paramagnetic compounds useful for MRI image enhancement can be conjugated to a substrate bearing paramagnetic ion chelators or exposed chelating functional groups, e.g., SH, NH_2, COOH, for the ions, or linkers for the radical addends. The foregoing are merely illustrative of the many methods of radiolabeling proteins known to the art. The MRI enhancing agent must be present in

sufficient amounts to enable detection by an external camera, using magnetic field strengths which are reasonably attainable and compatible with patient safety and instrumental design. The requirements for such agents are well known in the art for those agents which have their effect upon water molecules in the medium, and are disclosed inter alia, in, e.g., Pykett, Scientific American, 246:78 (1982); and Runge, *et al.*, Am. J. Radiol., 141:1209 (1987).

The following examples illustrate the manner in which the invention can be practiced. It is understood, however, that the examples are for the purpose of illustration and the invention is not to be regarded as limited to any of the specific materials or conditions therein.

Example 1

PCR Cloning of ZCE 025 Variable Regions

The initial cDNA cloning of the ZCE 025 Variable regions using the method of Okayama, H. and Berg, P. (Mol. and Cell. Biol., 2:161-170 (1982); Mol. and Cell. Biol., 3:280-289 (1983)) gave 3' sequences for both the heavy and light chains. In order to obtain the 5' sequences, the variable regions were isolated using a method termed "anchor PCR" (Loh, E.Y. et al.,Science, 243, 217-220 (1989)). Anchor PCR allows the use of a specific heavy or light chain primer (in our case, a sequence in the CK or CH1 regions) and a second poly-C-containing primer that recognizes a poly-G sequence added to all the mRNA-derived cDNAs, as is shown in **Table 1** below. Another advantage of this technique is that the upstream primer recognizes an added synthetic segment of DNA, making it possible to obtain the native sequence of the entire signal region.

Table 1

poly C primer C region primer poly A tail

CCCCCCC XXXXXXX

GGGGGGG----------XXXXXXX----------TTTTTTT

_____region amplified_____

a. Cloning of ZCE Kappa Light Chain cDNA

(1) ZCE 025 mRNA was obtained using the Guanidinium HCl procedure, as described in Sambrook, *et al.* (*supra*, 7.18-7.22).

(2) The first and second strand cDNA syntheses were performed using the Stratagene (San Diego, CA) LambdaZap® cDNA cloning kit according to the manufacturer's directions without the incorporation of a radioactive nucleotide. The resulting cDNA was ethanol precipitated.

(3) A poly G tail was added to the 3' ends of the cDNA by resuspending the precipitated cDNA in 23 µl water and adding 10 µL 5X tailing salts (0.9M Sodium Cacodylate, 150mM Tris-HCl (pH 6.8)), 5 µL 1mM Dithiothreitol, 5 µL 10 mM dGTP, 5 µl 10mM Cobalt Chloride, 2 µl (40 Units) terminal deoxynucleotide transferase (Boehringer Mannheim, Indianapolis, IN) and incubating for 1 hour at 37°.

(4) The poly G tailed cDNA was digested with Xho I. This enzyme cleaves the cDNA at an Xho I site within the Stratagene primer specific to the poly A region of the mRNA used for cDNA synthesis and removes the downstream poly G tail on the second strand of the cDNA.

(5) The ZCE 025 Kappa V region was isolated from the cDNA using the *Geneamp*® PCR kit from Perkin Elmer Cetus (Norwalk, CT) according to the manufacturer's instructions. The poly G-tailed, Xho I-cut cDNA was used as template with the following poly C upstream primer:

5'GAC TAG CGG CCG CAT CGA TCC CCC CCC CCC C (**SEQ. I.D. No. 3**) and a murine Kappa-specific downstream primer:

5'CAG ACG TCG ACG ATG GAT ACA GTT GGT GCA GCA TC (**SEQ. I.D. No. 4**) The amplification conditions were 94° for 1 min, 45° for 1 min, 72° for 3 min for 25 cycles.

(6) The amplified DNA was digested with Sal I and Not I and ligated into the pBluescript® cloning vector from Stratagene which vector had been previously digested with Sal I and Not I. The ligated mixture was used to transform freshly prepared competent cells of the *E. coli* strain MC1061 (Clonetech, Palo Alto, CA). The bacterial cells thus transformed were identified by ampicillin resistance.

(7) Positive colonies were confirmed by restriction enzyme analysis and these had inserts of approximately 400 bp, the expected size for the kappa V region.

(8) The positive clones were verified by sequence analysis on the Genesis® 2000 automated DNA sequencer from DuPont (Wilmington, DE). The cDNA sequence (**SEQ. I.D. NO. 5**) of the light chain variable region of ZCE 025 obtained and the corresponding amino acid sequence (**Sequence I.D. No. 6**)

SEQ. I.D. NO. 5

ZCE-025 Light Chain Variable cDNA

GAC ATT GTG ATG ACC CAG TCT CAA AAA TTT ATG TCC ACA TCA GTT GGA
GAC AGG GTC AAC ATC ACC TGC AAG GCC AGT CAG AAT GTT CGT ACT GCT
GTA GCC TGG TAT CAA CAG AAA CCA GGG CAG TCT CCT AAA GCA CTG ATT
TAC TTG GCA TCC AAC CGG TAC ACT GGA GTC CCT GAT CGC TTC ACA GGC
ATT GGA TCT GGG ACA GAT TTC ACG CTC ATC ATT AGC AAT GTG CAA TCT
GAA GAC CTG GCA GAT TAT TTC TGT CTG CAA CAT TGG AAT TAT CCT CTC
ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA C
381

SEQ. I.D. No. 6

Murine ZCE-025 Light Chain Variable Region Amino Acid Sequence:

DIVMTQSQKFMSTSVGDRVNITCKASQNVRTAVAWYQQKPGQSPKALIYLASNRYTGVPDR
FTGIGSGTDFTLIISNVQSEDLADYFCLQHWNYPLTFGAGTKLELK

b. Cloning of ZCE Gamma cDNA

(1) ZCE 025 mRNA was obtained using the Guanidinium HCl procedure, as described in Sambrook, *et al.*, (*supra*, 7.18-7.22).

(2) cDNA was prepared using the method described in **Example 1.a.**, above, for the ZCE kappa light chain.

(3) A poly G tail was added to the 3' ends of the cDNA as described in **Example 1.a.**, above.

(4) The poly G tailed cDNA was digested with Xho I and the ZCE 025 Gamma variable region was isolated from the cDNA using the *Geneamp*® PCR kit from Perkin Elmer Cetus according to the manufacturer's instructions. The poly G-tailed, Xho-I cut cDNA was used as template with the following poly C upstream primer:

5'GACTAGCGGCCGCATCGATCCCCCCCCCCCCCC (**SEQ. I.D. NO. 3**)

and a murine Gamma 1 specific downstream primer:

5'CAG ACG TCG ACG TTC CAG GTC ACT GTC ACT GGC TC (**SEQ. I.D. NO. 7**) The amplification conditions were 94° for 1 min, 45° for 1 min, 72° for 3 min for 40 cycles.

(6) The amplified DNA was digested with Sal I and Not I and ligated into the pBluescript® cloning vector (Stratagene, San Diego, CA) which had been previously digested with Sal I and Not I. The ligated mixture was used to transform freshly prepared competent cells of the *E. coli* strain MC1061. The bacterial cells thus transformed were identified by ampicillin resistance.

(7) Positive colonies were confirmed by restriction enzyme analysis and these had inserts of approximately 450 bp, the expected size for the Gamma chain variable region.

(8) The positive clones were verified by sequence analysis on the Genesis® 2000 automated DNA sequencer from DuPont (Wilmington, DE), according to the manufacturer's instructions. The cDNA sequence (**SEQ. I.D. NO. 8**) of the ZCE heavy chain variable region obtained and the corresponding amino acid sequence (**SEQ. I.D. NO. 9**)

SEQ. I.D. NO. 8

ZCE-025 Heavy Chain Variable cDNA Sequence:

GAT GTG CAG CTG GTG GAG TCT GGG GGA GGC TTA GTG CCG CCT GGA GGG
TCC CGG AAA CTC TCC TGT GCA GCC TCT GGA TTC ACT TTC AGT AAC TTT
GGA ATG CAC TGG ATT CGT CAG GCT CCA GAG AAG GGA CTG GAG TGG GTC
GCA TAC ATT AGT GGT GGC AGT AGT ACC GTC CAC TAT GCA GAC TCC TTG
AAG GGC CGA TTC ACC ATC TCC AGA GAC AAT CCC AAG AAC ACC CTG TTC
CTA CAA ATG ACC AGT CTA AGG TCT GAA GAC ACG GCC ATG TAT TAC TGT
GCA AGA GAT TAC TAC GTT AAT AAC TAC TGG TAC TTC GAT GTC TGG GGC
GCA GGG ACC ACG GTC ACC GTC TCC TCA G
420

SEQ. I.D. NO. 9

Murine ZCE-025 Heavy Chain Variable Region Amino Acid Sequence:

DVQLVESGGGLVPPGGSRKLSCAASGFTFSNFGMHWIRQAPEKGLEWVAYISGGSSTVHYA
DSLKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARDYYVNNYWYFDVWGAGTTVTVSS

Example 2

Cloning and Sequencing IM9 light and heavy chain cDNAs

The human **plasmacytoma** cell line IM9 (ATCC #159) expresses an IgG(γ_1, K) immunoglobulin.

a. Extraction of IM9 mRNA.

A total of $8X10^7$ IM9 cells were used for mRNA purification by the Fast-Trak™ kit from Invitrogen (San Diego, California) using an enzyme mix to digest the cells and oligo dT resin to adsorb the polyadenylated mRNA from the cell lysate according to manufacturer's directions. The resulting mRNA was redissolved in 100µl of sterile water and split into 10µl aliquots. Each aliquot was stored at -20° in ammonium acetate and ethanol.

b. Synthesis of cDNA.

The synthesis of a cDNA library was performed using a Librarian kit (Invitrogen). The pooled mRNA from four of the tubes in a. was quantitated by measuring absorbance at 260nm. The first strand cDNA synthesis was performed according to manufacturer's directions using an oligo-dT primer and reverse transcriptase in the presence of deoxyribonucleotides and RNAase inhibitors. Second strand synthesis was begun immediately by addition of ribonuclease H, *E. coli* ligase, and DNA polymerase in the presence of the appropriate buffer. The reaction was extracted once with phenol/chloroform and precipitated. The pellet was resuspended in sterile water and ligated with BstXI linkers supplied with the kit.

c. Purification of cDNAs.

The products of cDNA synthesis and linker ligation were separated by size on an agarose gel in TAE (tris acetate EDTA) buffer (see Sambrook, *et al., supra*). The cDNA molecules over 700bp were cut out of the gel and separated from the agarose by electroelution into a small volume of TAE buffer (0.04M Tris-acetate, 0.001M EDTA). The cDNA was extracted once with phenol/chloroform and precipitated. The sample was centrifuged, and the pellet was rinsed with ethanol, then air-dried.

d. Vector construction and transformation of bacterial cells.

The purified cDNA was ligated to the vector provided in the kit, pCDNAII, which is already cut with an enzyme that leaves the appropriate sticky ends for the linker used on the cDNA and not for relegation to itself. The ligation mixture was electroporated into the *E. coli* strain DH10B (ElectroMAX) (BRL, Gaithersburg, Maryland) using the Cell-Porator (BRL) at 330uF, 2.5kV. The total number of colonies obtained in this library was $1.8X10^6$ clones.

e. Preparation of Filter lifts of the IM9 cDNA library.

The library was inoculated onto LB agar media (950ml deionized water, bacto-tryptone 10g, bacto-yeast extract 5g, NaCl 10g) with ampicillin at 7500cfu (colony forming units) per 15cm plate. A total of 12 plates were made for a total of $9x10^4$cfu of cDNA clones. The colonies were blotted onto nylon filters by placing a dry filter on the colonies and removing the filter. The plates were returned to the incubator to allow the bacteria to grow back. The filters were placed on a layer of Whatman filter paper saturated with 5% SDS, 2 X SSC and put into the microwave oven on a high setting for 10 minutes. The filters were air-dried and stored at 4°C.

f. Primary Screening of the IM9 cDNA library.

The filters were incubated at 45°C in prehybridization buffer (2XSSC, 1%SDS, 0.5% nonfat dry milk). These were then hybridized with human Ig mixed kappa and gamma constant region probes using a method and probes described in C.B. Beidler, *et al.*, *supra*. The probes were labeled using a Prime-It® kit (BRL) in 6 X SSC, 1% SDS, 0.5% nonfat dry milk, at 65°C overnight. The filters were washed with 6 X SSC, 1% SDS, three times at 65°C, 5 minutes each time, then with 1 X SSC, 0.1% SDS, three times at 65°C, 20 minutes each time. The filters were put on Kodak XAR-5 X-ray film at **room temperature** overnight.

g. Secondary Screening of the IM9 cDNA library.

Sixty-two positive colonies were picked from the plates and streaked onto LB agar media in duplicate, twelve to a plate, for two sets of six plates. These were blotted on nylon filters and hybridized using a method and probes described in C.B. Beidler, *et al.*, *supra*. One set was hybridized with a kappa constant region probe and one with a gamma constant region probe.

h. Tertiary and Quaternary Screening of the IM9 cDNA library.

The streaks that were positive for the kappa probe were picked and plated out on LB media with ampicillin. These plates were blotted as described before, the filters were hybridized, and the positives were picked. These clones were subjected to one more round of blotting and hybridization to prove that the clone was pure. The sequence is provided below as Sequence I.D. 10 and the amino acid sequence is provided as Sequence I.D. 11.

SEQ. I.D. No. 10

GAC ATC CAG ATG ACC CAG TTT CCT TCC ACC CTG TCT GCT TCT GTA GGA GAC AGA GTC ACC 60
ATC ACT TGT CGG GCC AGT CAG AGT ATT AGT GCC TGG TTG GCC TGG TAT CAG CAG AAA CCA 120
GGG AAA GCC CCT AAA CTC CTG ATC TAT AAG GCG TCT AGT TTA GAA AGT GGG GTC CCA TCA 180
AGG TTC AGC GGC AGT GGA TCT GGG ACA GAG TTC ACT CTC ACC ATC ACC AGC CTG CAG CCT 240
GAT GAT TTT GCA ACT TAT TTC TGC CAA CAC TAT AAT CGA CCG TGG ACG TTC GGC CAA GGG 300
ACC AAG GTG GAA ATC AAA GCA

IM9 Light Protein SEQ I.D. No. 11

DIQMTQFPSTLSASVGDRVTITCRASQSISAWLAWYQQKPGKAPKLLIY
KASSLESGVPSRFSGSGSGTEFTLTITSLQPDDFATYFCQHYNRPWTFGQGTKVEIK

i. Southern blot and sequence analysis of light chain cDNA clones.

Ten putative kappa light chain clones were raised in LB broth with ampicillin. The plasmids were purified by the miniprep method of Holmes and Quigley (D.S. Holmes and M. Quigley, *Analytical Biochemistry*, 114:193, 1981). The

miniprep DNA was characterized by restriction enzyme mapping and Southern blot analysis. The longest of the cDNA inserts obtained (clone kappa LI) was 1.2kb. This clone was sequenced on the Genesis 2000 automated DNA sequencer (DuPont, Wilmington, Delaware) as described previously.

j. Rescreening, Southern blot, and sequence analysis of heavy chain cDNA clones.

None of the gamma clones was positive, so the library was rescreened as described in f. with the gamma constant region as a probe. The positives from this screening were picked and rescreened as described above in g. and h. until pure cultures were obtained. The putative clones were raised and characterized as described in i. and two gamma cDNA clones were found. The clones were both 1.6kb in length. The clones were sequenced on the Genesis 2000 automated DNA sequencer (DuPont) as described previously. The Sequence is provided below as Sequence I.D. 12 and the corresponding amino acid Sequence is provided as Sequence I.D. 13.

SEQ. I.D. No. 12

GAA ATG CAA CTG GTG GAA TTT GGG GGA GGC CTG CTA CAG CCT GGC AGG GCC CTG AGA CTC 60

TCC TGT GCA GCC TCT GGA TTC AGG TTT GAT GAT TAT GCC ATG CAC TGG GTC CGG CAA ACT 120

CCA GGG AAG GGC CTG GAG TGG GTC GCA GGT ATT AGT TGG AAT AGT GAC ACC ATA GAC TAT 180

GCG GAC TCT GTG AAG GGC CGA TTC ACC ATC TCC AGA GAC AAC GCC AAG AAC TCC CTC TAT 240

TTG CAA ATG AAC AGT CTC AGA GCT GAG GAC ACG GCC TTG TAT TAC TGT ACA AAA AGA AGG 300

GGG GTG ACA GAC ATT GAC CCT TTT GAT ATC TGG GGC CAA GGG ACA ATG GTC ATC GTC TCT 360

TCA GAG 366

IM9 HEAVY PROTEIN SEQ I.D. No. 13

EMQLVEFGGGLLQPGRALRLSCAASGFRFDDYAMHWVRQTPGKGLEWVAGISWNSDTIDYA
DSVKGRFTISRDNAKNSLYLQMNSLRAEDTALYYCTKRRGVTDIDPFDIWGQGTMVIVSS

Example 3

Definition of Structurally Conserved Regions and identification of chain association residues using known three-dimensional structures of antibodies.

a. Definition of Light Chain SCRs

First, the linear amino acid sequences of the light chain variable regions of a set of antibodies with known three-dimensional structures were compared. Eight sequences [Table 2] were compared in this example, but more or less may be used, by linear display of one sequence above the other on the computer screen [Figure 6] (SEQ. ID No. 14-21)

Table 2

Identifier	PDB File Name	Antibody Name	Source	Resolution
MCP	PDB1MCP.ENT	MCPC603	Mouse	2.7Å
FAB2	PDB4FAB.ENT	4-4-20	Mouse	2.7Å
HFL	PDB2HFL.ENT	HYHEL-5	Mouse	2.54Å
FDL	PDB1FDL.ENT	D1.3	Mouse	2.5Å
FBJ	PDB2FBJ.ENT	J539	Mouse	1.95Å
FAB1	PDB6FAB.ENT	36-71	Mouse	1.9Å
FAB	PDB3FAB.ENT	NEW	Human	2.0Å
FB4	PDB2FB4.ENT	KOL	Human	1.9Å

SEQ I.D. No 14

DIVMTQSPSSLSVSAGERVTMSCKSSQSLLNSGNQKNFLAWYQQKPGQPPK
LLIYGASTRESGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCQNDHSYPLTFGAGTKL

SEQ I.D. No 15

DVVMTQTPLSLPVSLGDQASISCRSSQSLVHSQGNTYLRWYLQKPGQSPKV
LIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTHVPWTFGGGTKLE

SEQ I.D. No 16

DIVLTQSPAIMSASPGEKVTMTCSASSSVNYMYWYQQKSGTSPKRWIYDTS
KLASGVPVRFSGSGSGTSYSLTISSMETEDAAEYYCQQWGRNPTFGGGTKLEIK

SEQ I.D. No 17

DIQMTQSPASLSASVGETVTITCRASGNIHNYLAWYQQKQGKSPQLLVYYT
TTLADGVPSRFSGSGSGTQYSLKINSLQPEDFGSYYCQHFWSTPRTFGGGTKLEIK

SEQ I.D. No 18

EIVLTQSPAITAASLGQKVTITCSASSSVSSLHWYQQKSGTSPKPWIYEIS
KLASGVPARFSGSGSGTSYSLTINTMEAEDAAIYYCQQWTYPLITFGAGTKLELK

SEQ I.D. No 19

DIQMTQIPSSLSASLGDRVSISCRASQDINNFLNWYQQKPDGTIKLLIYFT
SRSQSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGNALPRTFGGGTKLEIK

SEQ I.D. No 20

SVLTQPPSVSGAPGQRVTISCTGSSSNIGAGNHVKWYQQLPGTAPKLLIFH
NNARFSVSKSGSSATLAITGLQAEDEADYYCQSYDRSLRVFGGGTKLTVL

SEQ I.D. No 21

QSVLTQPPSASGTPGQRVTISCSGTSSNIGSSTVNWYQQLPGMAPKLLIYR
DAMRPSGVPDRFSGSKSGASASLAIGGLQSEDETDYYCAAWDVSLNAYVFGTGTKVTVL

Using the Insight II Homology software to facilitate the three-dimensional alignment of these structures, a landmark amino acid, known to be universally

conserved among antibodies, such as the cysteine at L23 (Kabat, E. A., et al., Sequences of Proteins of Immunological Interest, Vol 1 edition, U.S. Department of Health and Human Services, PHS, NIH, Bethesda, Maryland (1991)) was identified in each sequence. The sequences were vertically aligned on the computer screen.

Now, taking the first two of the linearly aligned sequences, one was designated to be held constant and the other to be superimposed onto the first (In practice, the bottom sequence on the display was held constant due to the program design.) A one residue box was drawn around the aligned cysteines. Then, using the commands for manual alignment of structures, the program determined the minimum RMS (Root Mean Square) deviation, after applying the optimum rotation and translation, of any boxed region. The minimum number of residues required in a box by this program before RMS deviation can be calculated in this way is three. As an integral part of this process a visual representation of the superimposed structures is displayed on the screen. A three residue box was made, using the program, centered on the residue of interest (here, the cysteine L23). The meaning of the box within this program is to mathematically superimpose the structures using the backbone atoms of the amino acids within the box. The box was moved horizontally one residue in each direction, sequentially. The position giving the lowest RMS deviation for the superposition of backbone atoms of the three amino acids from the linearly aligned sequences was selected.

The object of this preliminary step was to approximately superimpose the two structures, allowing structurally conserved regions (SCRs) to be discerned visually. Having achieved this objective, the box was now deleted. Using the already superimposed structures, SCRs [usually found in the regions of the beta sheets, but also in the other portions of the framework regions] were discovered by visual inspection. Using the Homology program, as described previously in this section, a box was made around

the amino acid sequences that the SCR comprises. Gaps were introduced in the structurally non-conserved (NSCR) regions to align the SCR sequences.

For each SCR defined in this way, the structures of each of the other six sequences were superimposed, sequentially, in each case holding the same sequence (first sequence) constant, and the appropriate boxes were determined. Then, the second SCR was identified for the initial two structures and the process was repeated working through each of the SCRs for all of the sequences (for example working from amino to carboxy terminus). Once the second set of SCRs was superimposed, the program was directed to superimpose the two structures based on all of the backbone atoms of the residues of both of the sets of SCRs. This process was also repeated for each of the subsequent sets of SCRs.

Now that SCR boxes had been determined for each of the sequences, consensus boxes were determined for each SCR. Consensus boxes represent the maximum number of amino acid positions (e.g. L60-L65 in **Figure 6**) contained in all of the SCR boxes at a particular site. In this example seven concensus SCR boxes were formed as shown in **Figure 6**.

b. Definition of heavy chain SCRs using known three-dimensional structures of antibodies.

First, the linear amino acid sequences of the heavy chain variable regions of a set of antibodies with known three-dimensional structures (we used eight sequences in this example, but more or less may be used) [**Table 2**] were compared by linear display of one sequence above the other on the computer screen [**Figure 7**] (**SEQ I.D. No. 22-29**)

SEQ I.D. No 22

EVKLVESGGGLVQPGGSLRLSCATSGFTFSDFYMEWVRQPPGKRLEWIAAS
RNKGNKYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCARNYYGSTWYFDVW
GAGTTVTVSS

SEQ I.D. No 23

EVKLDETGGGLVQPGRPMKLSCVASGFTFSDYWMNWVRQSPEKGLEWVAQIRNKPYNYETY
YSDSVKGRFTISRDDSKSSVYLQMNNLRVEDMGIYYCTGSYYGMDYWGQGTSVTVSS

SEQ I.D. No 24

VQLQQSGAELMKPGASVKISCKASGYTFSDYWIEWVKQRPGHGLEWIGEILPGSGSTNYHE
RFKGKATFTADTSSSTAYMQLNSLTSEDSGVYYCLHGNYDFDGWGQGTTLTVSS

SEQ I.D. No 25

QVQLKESGPGLVAPSQSLSITCTVSGFSLTGYGVNWVRQPPGKGLEWLGMIWGDGNTDYNS
ALKSRLSISKDNSKSQVFLKMNSLHTDDTARYYCARERDYRLDYWGQGTTLTVSS

SEQ I.D. No 26

EVKLLESGGGLVQPGGSLKLSCAASGFDFSKYWMSWVRQAPGKGLEWIGEIHPDSGTINYT
PSLKDKFIISRDNAKNSLYLQMSQVRSEDTALYYCARLHYYGYNAYWGQGTLVTVSA

SEQ I.D. No. 27

EVQLQQSGVELVRAGSSVKMSCKASGYTFTSNGINWVKQRPGQGLEWIGYNNPGNGYIAYN
EKFKGKTTLTVDKSSSTAYMQLRSLTSEDSAVYFCARSEYYGGSYKFDYWGQGTTLTVSS

SEQ I.D. No 28

VKLEQSGPGLVRPSQTLSLTCTVSGTSFDDYYSTWVRQPPGRGLEWIGYVFYHGTSDTDTP
LRSRVTMLVNTSKNQFSLRLSSVTAADTAVYYCARNLIAGCIDVWGQGSLVTVSS

SEQ I.D. No 29

EVQLVQSGGGVVQPGRSLRLSCSSSGFIFSSYAMYWVRQAPGKGLEWVAIIWDDGSDQHYA
DSVKGRFTISRNDSKNTLFLQMDSLRPEDTGVYFCARDGGHGFCSSASCFGPDYWGQGTPV
TVSS

Using the Insight II homology software to facilitate the three-dimensional alignment of these structures, a landmark amino acid, known to be universally conserved among antibodies, such as the cysteine at H22 (Kabat, E. A., et al., supra) was identified in each sequence. The sequences were vertically aligned on the computer screen.

Now, taking the first two of the linearly aligned sequences one was designated to be held constant and the other to be superimposed onto the first. (In practice, the bottom sequence on the display was held constant due to the program design.) A one residue box was drawn around the aligned cysteines. Then, using the commands for manual alignment of structures, the program determined the minimum RMS deviation, after applying the optimum rotation and translation, of any boxed region. The minimum number of residues required in a box by this program before RMS deviation can be calculated in this way is three. As an integral part of this process, a visual representation of the superimposed structures is displayed on the screen. A three residue box was made, using the program, centered on the residue of interest (here, the cysteine H22). The meaning of the box within this program is to mathematically superimpose the structures using the backbone atoms of the amino acids within the box. The box was moved horizontally one residue in each direction, sequentially, and the position giving the lowest RMS deviation for the superposition of backbone atoms of the three amino acids from the linearly aligned sequences was selected.

The object of this preliminary step was to approximately superimpose the two structures, allowing SCRs to be discerned visually. Having achieved this objective, the box was now deleted. Using the already superimposed structures, SCRs [usually found in the regions of the beta sheets, but also in other portions of the framework regions] are discovered by visual inspection and put within boxes including appropriate amino acids, guided by the RMS deviations. Once vertically aligned, the box was expanded in

both directions to include more amino acids, until the RMS deviation became unacceptable (usually >0.75 Å). Then the size of the box was reduced to the size which had the last acceptable RMS deviation. Gaps are introduced in the structurally non-conserved (non-homologous) regions to help align the SCRs vertically.

For each SCR defined in this way, the structures of each of the other six sequences were superimposed, sequentially, in each case holding the same sequence (first sequence) constant, and the appropriate boxes were determined. Then, the next SCR was identified for the initial two structures and the process was repeated working through each of the SCRs for all of the sequences (for example working from amino to carboxy terminus). Once the second set of SCRs was superimposed, the program was directed to superimpose the two structures based on all of the backbone atoms of the residues of both of the sets of SCRs. This process was also repeated for each of the subsequent sets of SCRs.

Now that SCR boxes had been determined for each of the sequences, consensus boxes were determined for each SCR. Consensus boxes represent the maximum number of amino acid positions (e.g. H3-H6 of antibody FB4 in **Figure 7**) contained in all of the SCR boxes at a particular site. Thus, the amino acids contained in each consensus SCR box are structurally conserved among all of the database antibodies under consideration. In this example ten concensus SCR boxes were formed as shown in **Figure 7**.

c. Identification of Chain Association residues in known structures.

For each of the known structures used in defining the SCRs, described above, chain-association residues were identified. First, all residues from the light chain which contain any atom which is within about 4.5 Å of any atom of any heavy chain residue, except those of the Kabat-defined CDRs, were identified. This set was then limited to those

which by orientation have a significant likelihood of interaction with any atom of that heavy chain residue. All residues from the heavy chain which contain any atom which is within about 4.5Å of any atom of any light chain residue, except those of the Kabat-defined CDRs, were also identified, again limited to those which have a significant likelihood of interaction. This process was carried out for each of the antibodies of known structure shown in **Table 2**.

Example 4

Three-dimensional Modeling of ZCE Fv

a. Three-dimensional modeling of ZCE Light chain variable domain.

The three dimensional coordinates had not been determined for ZCE Fv. For this reason, homology modeling was used to approximate the actual structure. The following four steps were used: (1) alignment of the ZCE light chain variable region sequence with the aligned sequences of the set of light chain variable regions of known structure described in **Example 3.a.**; (2) homology modeling of SCRs using SCRs from the known light chain variable region structures; (3) homology modeling of NSCRs using the full range of known structures available in the Brookhaven database, and (4) a series of energy minimizations carried out to obtain an energetically favorable structure.

(1) Alignment of ZCE 025 light chain amino acid sequence with amino acid sequences of known light chain structures. The linear sequence of the ZCE 025 light chain variable region, determined from a cDNA clone as described in **Example 1.a.**, above, was displayed and aligned with the database sequences described in step 1 above, using the Insight II software. As described for the database sequences, the first step was to align the ZCE 025 sequence with the database sequences using the first consensus SCR box. This was accomplished by first

identifying which residues (one or more) within the box were most highly conserved between the known structures, identifying these residues in the ZCE sequences, and aligning them. In some cases only a subset of the residues identified as conserved appeared in the ZCE sequence. In these instances, the subset was aligned. Working from one concensus SCR box to the next (in this example, we worked from amino to carboxy terminus) the process was repeated. Where necessary, gaps were introduced either into regions other than those corresponding to SCRs (i.e. NSCRs) from ZCE or into identical positions within the SCRs of each of the aligned known structures.

The object of this preliminary step was to align the ZCE sequence with the sequences of the other light chain variable regions of known structure. In each case great effort was made to identify the potential locations of ZCE SCRs by linear sequence homology to the consensus regions alone. The result of this alignment is shown in **Figure 8**.

(2) - Three dimensional modeling of SCRs. For each SCR, the actual known structure whose sequence has the greatest homology to the corresponding ZCE light chain SCR was selected as template for that segment, and its coordinates were assigned to the ZCE SCR. If there were a residue in a template SCR that did not match the corresponding residue in the ZCE SCR residue, the residue in the template was mutated to match the ZCE SCR residue, while maintaining the coordinates of all the atoms in the backbone and side chains of the template residue that correspond to those in the ZCE residue and modeling the remaining atoms under the constraints of maintaining the same bond lengths, angles and dihedrals as those in the original database residue, e.g., for gamma and delta carbons. This was done for each SCR (we worked from amino to carboxy terminus). After all of the SCRs were assigned coordinates in this manner a partial three-dimensional structure comprising the modeled SCRs was displayed, absent the NSCRs.

(3) - Three dimensional modeling of NSCRs. For each ZCE light chain NSCR, the flanking SCRs of assigned coordinates were used along with the length of the NSCR to identify a known structure with the greatest likelihood of being homologous to the SCR/NSCR/SCR array. This was accomplished by using the "Loop Search" subprogram in Insight II to search the database for structures (1) containing proper lengths of flanking and spanning sequences and (2) having coordinates for the flanking sequences with the least RMS deviation from those of the assigned SCRs. In practice, approximately ten structures (more or less can be used) were ranked by the program on the basis of RMS deviation of the flanking sequences. These were sequentially displayed on the screen superimposed on the flanking SCRs. The structure best approximating the flanking sequences and having the same general orientation as NSCRs from light chain variable regions of known structure was chosen as template for that particular NSCR and its coordinates were assigned to the ZCE NSCR. This process was then repeated for each NSCR, and the NSCRs were added to the computer model by inserting each in its appropriate place, for instance flanked by the adjoining SCRs, until the entire variable region had been modeled.

(4) - Energy Minimizations of modeled structure. Energy minimizations were carried out in stages to assure that no major structural disruptions would occur. Once all of the NSCRs making up the model had in turn been selected from the database, fixed in space, and modeled to transform them into the corresponding ZCE NSCRs, the splice regions where the SCRs join the NSCRs were refined to relieve any strain in the model that would result from joining the SCRs and NSCRs, using the "Repair" algorithm to assign the proper bond lengths, bond angles, and omega values to the structures.

Now, the "Relax" algorithm was applied in a series of sequential steps to the model as a whole: (1) to the side chains of the NSCRs to assign proper geometries, and remove

any unfavorable non-bonded contacts between side chain atoms and other atoms in the molecule, (2) to all atoms of the NSCRs, and (3) to the mutated side chains of the SCRs. In each of these steps, all regions other than those that are being relaxed remain fixed to their assigned coordinates.

Finally, an energy minimization analysis was performed using the "Discover" subprogram to allow the model to assume an energetically favorable structure. First the entire model was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/$Å^2$). Then energy minimization was performed on the entire molecule without the backbone atoms being tethered.

The result of carrying out these steps was a model of the ZCE light chain.

b. Three-dimensional modeling of ZCE 025 Heavy chain variable domain.

Like the light chain, coordinates had not been determined for ZCE heavy chain. For this reason, homology modeling was again used to approximate the actual structure. The same steps were used for heavy chain as for light.

(1) - Alignment of ZCE 025 heavy chain amino acid sequence with amino acid sequences of known heavy chain structures. The linear sequence of the ZCE heavy chain variable region, determined from a cDNA clone as described in **Example 1.b.**, above, was displayed and aligned with the database sequences described in **Example 3.b.** above, using the Insight II software. As described for the light chain, the first step was to align the ZCE sequence with the database sequences using the first consensus SCR box. The remainder of the process was precisely as described for the light chain, with the final alignment displayed in **Figure 9**.

(2) - Three dimensional modeling of SCRs. For each SCR, the actual known structure whose sequence has the greatest

homology to the corresponding ZCE SCR was selected as template for that particular SCR (these are shown in bold in **Figure 9** and its coordinates were assigned to the ZCE SCR. This was done for each SCR (we worked from amino to carboxy terminus). After all of the SCRs were assigned coordinates in this manner a partial three-dimensional structure was displayed, absent the NSCRs.

(3) - Three dimensional modeling of NSCRs. As for light chain, for each ZCE heavy chain NSCR, the flanking SCRs which had been assigned coordinates were used along with the length of the NSCR to identify a known structure with the greatest likelihood of being homologous to the SCR/NSCR/SCR array, using the "Loop Search" subprogram in Insight II. This process was then repeated for each NSCR, until the entire variable region had been modeled.

(4) - Energy Minimizations of modeled structure. Energy minimizations were carried out in stages, as for the light chain, to assure that no major structural disruptions would occur. The process was identical to that described for the light chain, including (1) use of the "Repair" algorithm to assign the proper bond lengths, bond angles, and omega values to the structures; (2) use of the "Relax" algorithm to assign proper geometries and remove any unfavorable non-bonded contacts; (3) use of the "Discover" subprogram to allow the model to assume an energetically favorable structure. Once again, the entire molecule was first subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/$Å^2$). Then energy minimization was performed on the entire molecule without the backbone atoms being tethered.

The result of carrying out these steps was a model of the ZCE heavy chain.

c. Three-dimensional modeling of ZCE 025 heavy and light chains together to form Fv.

Now that coordinates had been determined for both light and heavy chain, these were displayed on the screen together. Since energy minimization has been carried out on each chain separately, the first step was to carry out the same procedure on the chains as a set. Finally, an energy minimization was performed using the "Discover" subprogram to allow the model to assume an energetically favorable configuration. First, potential chain association residues for ZCE light and heavy chains were identified by comparison with the chain association residues of the known structures (determined in **Example 3.c.**). Chain association residues in the aligned sequences were compared with residues in the corresponding position in ZCE. When an identical residue was present, it was designated as favorable to chain association; if a different residue was found, it was designated as potentially disrupting to chain association. Totals of favorable and disrupting residues were determined for the comparison of ZCE light and heavy chains to each of the known structures. The known structure providing the comparison having the greatest excess of favorable residues over disruptive residues was chosen as template for ZCE heavy/light association. If two or more known structures had the same excess of favorable over disruptive residues, the structure having the greatest number of favorable residues over disruptive residues was chosen. In this example 2HFL was chosen.

Next, the light chain structure determined for ZCE in **Example 4.a.** was superimposed on that of the light chain structure of 2HFL, using the backbone coordinates of the favorable residues described above. This was carried out using the "superimpose" command in the Insight II software. The same was done for the ZCE heavy chain using the 2HFL heavy chain. Next the entire molecule was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually

100Kcal/Å^2). Then energy minimization was performed on the entire light/heavy association model without the backbone atoms being tethered.

d. Identification of CDR-Associated Residues in ZCE 025 Fv.

For ZCE, the aim of modeling is to identify regions that must be conserved to conserve the function of the CDRs. To do this it is necessary to (1) identify all potential CDR-associated residues and (2) identify the subset of these which have a reasonable likelihood of a significant interaction with the CDR residue involved.

To determine which amino acids lying outside of the Kabat-defined CDRs may influence binding of the CDRs, all the amino acid residues outside of the Kabat-defined CDR regions that have an atom located within about 4.5 Angstroms of any atom in an amino acid located in the Kabat-defined CDR regions of the ZCE construct were identified as CDR-associated residues. As these were predicted to be important for maintaining the binding specificity of the ZCE antibody, they were earmarked for preservation as donor amino acids in the CDR grafted antibody construct in addition to those in the defined CDR regions.

We first identified all residues on the light or heavy chain that have atoms that are within 4.5 Å of any atoms of any light chain CDR residue. The set was limited to those with a significant likelihood of interaction, based on orientation of the residue, charge, hydrophobicity, etc. Next, all residues on the light or heavy chain that contain atoms which are within 4.5 Å of any atom of any heavy chain CDR residue were identified. Again, the set was limited to those with a high likelihood of significant interaction with the CDR residue of interest. In this way, the entire set of light and heavy chain CDR-associated residues was determined.

Example 5
Three dimensional modeling of IM9 Fv.

a. Three-dimensional modeling of IM9 Light chain variable domain.

Homology modeling was used to approximate the structure of the IM9 antibody. This process had four steps: (1) alignment of the IM9 light chain variable region sequence with the aligned sequences of the set of light chain variable regions of known structure. (See **Example 3.a.** above); (2) homology modeling of the IM9 light chain SCRs using SCRs from the known light chain variable region structures; (3) homology modeling of NSCRs (non-structurally conserved regions) using the full range of known structures available in the Brookhaven database (other known structures could also be used); and (4) a series of energy minimization routines to determine the energetically preferred structure.

(1) Alignment of IM9 light chain amino acid sequence with amino acid sequences of known light chains.

The linear DNA sequence [**SEQ. I.D. No. 10**] of the IM9 light chain variable region was determined from a cDNA clone as described in **Example 2.i.**:

The linear amino acid sequence [**SEQ. I.D. No. 11**] of the IM9 light chain variable domain was displayed on the computer screen and aligned with the sequences of the eight light chain variable regions of known structure described in **Example 3.a.** above, using the Insight II software. The IM9 sequence was aligned with the database sequences using the first consensus SCR box. The residues (one or more) within the box which were most highly conserved between the known structures were identified, after which the corresponding residues in the IM9 sequences were identified and the structures were aligned. When a subset of the residues identified as conserved in the known structures appeared in the IM9 sequence the subset was aligned. The alignment proceeded from one consensus SCR box to the next as described

above for the known sequences. The alignment proceeded from amino to carboxy terminus, but would work as well if reversed. Gaps were introduced either into regions other than those corresponding to SCRs (i.e. NSCRs) from IM9 or into identical positions within the SCRs of each of the aligned known structures when necessary for alignment.

This preliminary step allowed alignment of the IM9 sequence with the sequences of the other light chain variable regions of known structure. The potential locations of IM9 SCRs were identified by linear sequence homology to the consensus regions. This alignment is shown in **Figure 10**.

(2) Three-dimensional modeling of IM9 light chain SCRs.

For each SCR, the actual known light chain structure whose sequence had the greatest homology to the corresponding IM9 light chain SCR was selected as the template for that segment (these are shown in bold in **Figure 10** and its coordinates assigned to the IM9 SCR. In instances where a residue in a template SCR did not match the corresponding residue in the IM9 SCR, the coordinates of all the atoms in the backbone and sidechains of the template residue that correspond to those in the IM9 residue were maintained. The remaining atoms (e.g., for gamma and delta carbons and the atoms bonded to them) were modeled under the constraints of maintaining the same bond lengths, angles and dihedrals as those in the original database residue. This was done for each SCR (we worked from amino to carboxy terminus). After all of the SCRs were assigned coordinates in this manner a partial three-dimensional structure comprising the modeled SCRs was displayed, absent the NSCRs.

(3) Three-dimensional modeling of IM9 light chain NSCRs.

For each IM9 light chain NSCR, the flanking SCRs which had been assigned coordinates were used along with the length of the NSCR to identify a known structure with the

greatest likelihood of being structurally homologous to the SCR components of the SCR/NSCR/SCR array. In addition, the known structure containing a region corresponding to the NSCR component of the aforementioned SCR/NSCR/SCR array, is identified which has an orientation most like that of the corresponding region of the antibodies of known structure. This was accomplished by using the "Loop Search" subprogram in Insight II to search the database for structures (1) containing proper lengths of flanking and spanning sequences and (2) having backbone coordinates for the flanking sequences with the least RMS deviation from those of the assigned SCRs. In practice, a maximum of ten structures (more or less can be used depending on the limitations of the program used) were ranked by the program on the basis of RMS deviation of the coordinates of the backbone atoms of the flanking sequences. These were sequentially displayed on the screen superimposed on the flanking SCRs. The structure best approximating that of the flanking sequences, having the same general orientation as NSCRs from light chain variable regions of known structure, and having a minimum of structurally significant mutations was chosen as template for that particular NSCR and its coordinates were assigned to the NSCR. This process was then repeated for each NSCR, until the entire variable region had been modeled.

(4) Energy Minimizations of modeled IM9 light chain structure.

Energy minimizations were carried out in stages to assure that no major structural disruptions would occur. First the splice regions where the SCRs join the NSCRs were refined to relieve any strain in the model that would result from joining the SCRs and NSCRs, using the "Repair" algorithm to assign the proper bond lengths, bond angles, and omega values to the residues in the splice region.

Then, the "Relax" algorithm was sequentially applied to the regions as follows: (1) to the sidechains of the NSCRs to assign proper geometries, and remove any unfavorable non-bonded contacts between NSCR sidechain atoms and other atoms in the molecule; (2) to all atoms of the NSCRs to remove remaining unfavorable contacts between the NSCR and other atoms in the molecule; (3) to the altered side chains of the SCRs to remove any unfavorable non-bonded contacts between mutated SCR side chain atoms and other atoms in the molecule, and (4) to all the sidechain atoms of the SCR to remove remaining unfavorable side chain contacts. In each of the above described steps, all regions other than those which are being "relaxed" remain fixed to their assigned coordinates.

Finally, an energy minimization was performed using the "Discover" program to allow the model to assume an energetically favorable structure. First the entire model was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/$Å^2$). Then energy minimization was performed on the entire molecule without the backbone atoms being tethered.

The result of carrying out these steps was the homology model of the IM9 light chain.

b. Three-dimensional modeling of IM9 Heavy chain variable domain.

The steps used to model the IM9 heavy chain are similar to those used in modeling the IM9 light chain.

(1) Alignment of IM9 heavy chain amino acid sequence with amino acid sequences of known heavy chain structures.

The linear DNA sequence [**SEQ. I.D. NO. 12**] of the IM9 heavy chain variable region was determined from a cDNA clone as described in **Example 2.j.**

The linear amino acid sequence [**SEQ. I.D. NO. 13**] of the IM9 heavy chain variable domain was displayed and aligned with the database sequences, described in **Example 3.b.** above, using the Insight II software. As described for the light chain, the first step was to align the IM9 sequence with the database sequences using the first consensus SCR box. The remainder of the process was precisely as described for the light chain, with the final alignment displayed in **Figure 11.**

(2) Three-dimensional modeling of IM9 heavy chain SCRs.

For each SCR, the actual known structure whose sequence has the greatest homology to the corresponding IM9 SCR was selected as the template (shown in bold in **Figure 11**) and its coordinates assigned to the corresponding IM9 SCR. The process - working from amino to carboxy terminus - was repeated for each SCR. After all of the SCRs were assigned coordinates a partial three-dimensional structure was displayed, absent the NSCRs.

(3) Three-dimensional modeling of IM9 heavy chain NSCRs.

As for the light chain, for each IM9 heavy chain NSCR, the flanking SCRs which had been assigned coordinates

were used along with the length of the NSCR to identify a known structure with the greatest likelihood of being structurally homologous to the SCR components of the SCR/NSCR/SCR array.

In addition, the known structure containing a region corresponding to the NSCR component of the aforementioned SCR/NSCR/SCR array, is identified which has an orientation most like that of the corresponding region of the antibodies of known structure. This was accomplished by using the "Loop Search" subprogram in Insight II to search the database. This process was then repeated for each NSCR, until the entire variable region had been modeled.

(4) Energy Minimizations of modeled IM9 heavy chain structure.

Energy minimizations were carried out in stages, as for the light chain, to assure that no major structural disruptions would occur. The process used was in substantial accordance with that described for the light chain. The process comprised the following steps: (1) use of the "Repair" algorithm to assign the proper bond lengths, bond angles, and omega values to the splice regions; (2) use of the "Relax" algorithm to assign proper geometries and remove any unfavorable non-bonded contacts from the mathematical model; (3) use of the "Discover" subprogram to allow the model to assume an energetically favorable structure. As described for the light chain, the entire molecule was first subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/$Å^2$). Then energy minimization was performed on the entire molecule without the backbone atoms being tethered.

The resultant structure was used as the model of the IM9 heavy chain.

c. Three-dimensional modeling of IM9 heavy and light chains together to form Fv.

The coordinates determined for the light and heavy chain, were used to generate a model of the Fv. As an initial step, potential chain association residues for IM9 light and heavy chains were identified by comparison with the chain association residues of the known structures (determined in **Example 3.c.**). Chain association residues in the aligned sequences were compared with residues in the corresponding position in IM9. When an identical residue was present, it was designated as favorable to chain association; if a different residue was found, it was designated as potentially disrupting to chain association. Totals of favorable and disrupting residues were determined for the comparison of IM9 light and heavy chains to each of the known structures. The known structure providing the comparison having the greatest excess of favorable residues over disruptive residues was chosen as template for IM9 heavy/light association. If two or more known structures had the same excess of favorable over disruptive residues, the structure having the greatest number of favorable residues was chosen as template. In this example, FDL was chosen.

Next, the light chain structure determined for IM9 in **Example 5.a.** was superimposed on the template light chain structure of FDL, using the backbone coordinates of the favorable residues described above. This was carried out using the "superimpose" command in the Insight II software. The same was done for the IM9 heavy chain using the FDL heavy chain. Next the entire molecule was subjected to an energy minimization with the backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/$Å^2$). Then an energy minimization was performed on the entire light/heavy associated (Fv) model without the backbone atoms being tethered.

d. Identification of Chain-Association Residues in IM9 Fv.

The regions of IM9 that should be conserved to allow for optimal associations between the chains in regions other than those that will be replaced (the CDRs and CDR associated regions) was determined by (1) identification of all chain association residues; (2) identification of all CDR associated residues; and (3) delineation of the not CDR-associated subset of chain association residues. The individual steps are described in detail below.

Residues from the light chain that contain an atom that is within about 4.5Å of any atom of any heavy chain residue were identified. This set was then limited to those residues that have a significant likelihood of interacting with that heavy chain residue (or any other). All residues from the heavy chain containing an atom that is within about 4.5 Å of any atom of any light chain residue were identified, again limited to those that have a significant likelihood of interaction. Next, all residues on the light or heavy chain that contain an atom that is within about 4.5Å of any atom of any light chain CDR residue were identified. Again the set is limited to those with a significant likelihood of interaction. Next, all residues on the light or heavy chain that contain an atom that is within about 4.5 Å of any atom of any heavy chain CDR residue with a high likelihood of significant interaction with the CDR residue of interest were identified. Finally, the subset of chain association residues not contained within either set of CDR-associated residues was determined and classed as IM9 chain-association residues.

Example 6
Three-dimensional modeling of Humanized ZCE 025 Fv.

a. Modeling of CDR-grafted ZCE/IM9 light chain variable region.

The IM9 and ZCE light chain amino acid sequences were aligned with reference to the sequences of the eight known structures. On the ZCE linear array, the Kabat-defined CDRs and the CDR-associated residues determined in **Example 4** were identified. For SCR or NSCR regions which do not contain a CDR or CDR-associated residue, the entire region was replaced with the IM9 sequence. For SCRs which contain one or more CDR or CDR-associated residues, the non-CDR and non-CDR-associated residues were replaced with IM9 sequence, but the ZCE sequence was conserved for the CDR or CDR-associated residues. For NSCRs which contain one or more CDR or CDR-associated residues, the replacement is dependent upon the relative lengths of the region of interest in acceptor and donor molecules. If the NSCR has the same number of residues in both the acceptor (IM9) and the donor (ZCE) molecules, the non-CDR associated residues were replaced with acceptor (IM9) sequence. If however, the NSCR differs in number of residues between the acceptor and donor, the donor (ZCE) sequence was conserved for the entire segment. In this way the primary sequence for the light chain CDR-grafted molecule was determined. The residues of the CDR-grafted primary sequence were assigned coordinates to match those of the residues in the light chain sequences of the superimposed models of ZCE and IM9 from which they were derived. This was done working from amino to carboxy terminus.

b. Modeling of CDR-grafted ZCE/IM9 heavy chain variable region.

The IM9 and ZCE heavy chain amino acid sequences were aligned with reference to the sequences of the eight known heavy chain structures. On the ZCE linear array, the Kabat-defined CDRs and the CDR-associated residues determined

in **Example 4** were identified. For SCR or NSCR regions which do not contain a CDR or CDR-associated residue, the entire region was replaced with the IM9 sequence. For SCRs which contain one or more CDR or CDR-associated residues, the non-CDR and non-CDR-associated residues were replaced with IM9 sequence, but ZCE sequence was conserved for the CDR or CDR-associated residues. For NSCRs which contain one or more CDR or CDR-associated residues, the ZCE sequence was conserved for the entire region. In this way the amino acid sequence for the heavy chain CDR-grafted molecule was determined. The coordinates of the residues of the CDR-grafted primary sequence were obtained from those of the residues in the heavy chain sequences of the superimposed models of ZCE and IM9 from which they were derived. This was done working from amino to carboxy terminus.

c. Modeling of Humanized ZCE Fv.

Now that coordinates had been assigned for both light and heavy chain, these were displayed on the screen together. An energy minimization was performed using the "Discover" subprogram to allow the model to assume an energetically favorable structure. First the entire model was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a defined force constant (usually 100 Kcal/Å^2). Then the energy minimization was performed on the entire model without the backbone atoms being tethered.

d. Modification of the humanized ZCE 025 model so that only CDR-associated residues found in the murine ZCE 025 model meet the definition of CDR-associated residues.

CDR-associated residues were determined for the modeled humanized ZCE Fv in substantial accordance with the methodology taught for the original ZCE Fv. First, all residues on the light or heavy chain which contain atoms which are within about 4.5 Å of any atoms of any light chain

CDR residue, which also has a significant likelihood of interaction, based on orientation, charge, hydrophobicity, etc. were identified. Next, all atoms of all residues on the light or heavy chain which are within about 4.5 Å of any atoms of any heavy chain CDR residue were identified. The set was then limited to those with a high likelihood of significant interaction with any atoms of the CDR residue of interest. The entire set of light and heavy chain CDR-associated residues was thusly determined.

The set of CDR-associated residues determined for the humanized Fv was compared to that determined for the ZCE Fv. In any case where an additional CDR-associated residue was present in the humanized, the amino acid at that position was replaced by the amino acid found in the murine ZCE. In the case where a CDR-associated residue in ZCE was not identified as CDR-associated in the humanized ZCE and is found in a NSCR, the entire NSCR was changed to the donor (ZCE) sequence.

e. Confirmation of Chain-association residues.

After the CDR-associated residues were modified if necessary as described above, the model was analyzed to determine if the chain association residues identified for IM9 were conserved. In this example, they were conserved. If, however, differences are observed, they are noted, but no changes are made at this time. If, in addition, a significant decrease in secreted protein is observed for the humanized molecule, these are potential sites for modification.

The amino acid sequences for light and heavy chain hZCE, determined above, are shown in **Figure 7** and **Figure 8**, respectively.

Example 7

Modeling of hZCE-CSVL and hZCE-kb Fv.

a. Modeling of hZCE-CSVL.

The IM9 light and ZCE heavy chain primary amino acid sequences had already been aligned with reference to different sequences. Therefore, it was necessary to bridge these alignments through realignment using a common sequence. The IM9 heavy chain sequence was used for this purpose as shown in **Figure 12**. In addition, the IM9 heavy chain provided information on chain association residues. ZCE heavy chain sequence was added and aligned with the linear array containing light chain ZCE and light and heavy chain IM9 sequences. Once aligned in this manner, SCRs were defined there between as described in **Example 3**, the Kabat defined CDRs and CDR-associated residues determined in **Example 4**, were identified on the ZCE heavy chain linear array. For SCR or NSCR regions which do not contain a CDR or CDR-associated residue, the entire region was replaced with the IM9 light chain sequence (and structure, i.e., coordinates). For SCRs which contain one or more CDR or CDR-associated regions, the non-CDR-associated residues were replaced with IM9 sequence (and structure, i.e., coordinates), but ZCE heavy chain sequence (and structure, i.e., coordinates) was conserved for the CDR-associated residues. For NSCRs that contain one or more CDR or CDR-associated residues, the ZCE heavy chain sequence (and structure, i.e., coordinates) was conserved for the entire region. In this way the primary sequence for the heavy chain CDR-grafted molecule was determined, and a composite structure was developed.

Now, the resultant model was modified to assure that chain association residues, derived from the IM9 model were conserved. In all non-CDR or non-CDR-associated regions, when the amino acid in the position occupied by the chain association residue was different than the

corresponding IM9 heavy chain residue, it was replaced with the corresponding IM9 heavy chain chain-association residue. In this example no chain-association residues were found to lie in the CDR or CDR-associated regions. In the unlikely event that this should occur, the residue should be noted, but no change should be made. In addition one residue leucine at position 94, of the mature CSVL was changed to a methionine. The final amino acid sequence of the mature CSVL is shown in **Figure 13 [SEQ I.D. No. 30]**

SEQ I.D. No. 30

DIQMTQFPST LSASVGDRVN ITCRASGFTF SNFGMHWIRQ KPGKGLKWVA
YISGGSSTVH YADSLKGRFT ISRDNPKNEL FLTITSLQPD DFAMYYCARD
YYVNNYWYFD VWGQGTKVEI KR (122 residues)

Alternatively, hZCE light chain can be used as acceptor and hZCE heavy chain can be used as donor. In this case, chain association residues used for the preliminary Fv model are those identified for hZCE FV.

b. Model hZCE-kb Fv.

Now that coordinates had been assigned for both light and heavy/light hybrid chains, these were displayed on the screen together. An energy minimization was performed using the "Discover" subprogram to allow the model to assume an energetically favorable configuration. First the entire model was subjected to energy minimization with backbone atoms tethered to their starting coordinates with a force constant of 100 Kcal/Å^2. Then the energy minimization algorithm was applied to the entire model without the backbone atoms being tethered.

c. Modify to assure no added CDR-associated residues.

CDR-Associated residues were determined for the modeled humanized ZCE light chain dimer as for the original ZCE Fv of **Example 4.d.** Again, this was done by first

identifying all residues on the light or heavy/light hybrid chain that are within 4.5 Å of any light chain CDR residue, and that also have a significant likelihood of interaction, based on orientation of the residue, charge, hydrophobicity, etc. Next, all residues on the light or heavy/light hybrid chain that were within 4.5 Å of any heavy/light hybrid chain CDR residue were identified. Again, the set was limited to those with a high likelihood of significant interaction with the CDR residue of interest. In this way; the entire set of light and heavy/light hybrid chain CDR-associated residues was determined.

The set of CDR-associated residues determined for the humanized light chain dimer was compared to that determined for the ZCE Fv. In any case where an additional CDR-associated residue is seen for the humanized, the amino acid at that position was replaced by the amino acid found in the murine ZCE. Care should be taken in this step as these replacements would be dependent upon whether that residue lies in an SCR or NSCR segment as explained in **Example 6** above.

d. Confirm Chain-association residues.

After the CDR-associated residues were modified as necessary, the model was analyzed to determine if the chain association residues identified for IM9 were conserved. In this example, they were conserved. If, however, differences are observed, these are noted, but no changes are made at this time; If there is a significant decrease in expression observed for the humanized molecule, these are potential sites for modification.

Example 8

Construction of expression vectors pGIM9K and pGIM9K/hZCR-kappa.

a. **Construction and Screening the IM9 Genomic library in E. coli Bacteriophage Lambda for the Ig Kappa Gene.**

IM9 genomic DNA was extracted and purified using methods described in Sambrook (supra, pp. 9.4-9.30). The DNA was partially digested with MboI and separated by sucrose density gradient ultra-centrifugation. The gradients were fractionated and the aliquots were analyzed for size by agarose gel electrophoresis, as described in Sambrook (supra, pp. 6.3-6.19). The fractions between 8-20 Kb were pooled, and dialyzed against TE Buffer (10 mM Tris HCl; 1 mM EDTA, pH 7.4). "Tris" is [Tris(hydroxymethyl)amino methane].

The IM9 DNA was ligated to Lambda EMBL3 arms (commercially available from Stratagene, San Diego, California) and packaged with the lambda bacteriophage packaging kit, Gigapack® Gold (Stratagene). The recombinant bacteriophage particles were used to transfect *E. coli* strain P2/392, which was inoculated onto 1% NZY agar medium in 140 mm diameter plates. The lambda library contained 6.55 X 10^5 individual clones, and was amplified by plating at 3.3 X 10^4 plaques per plate on twenty plates and suspending the bacteriophage in 200 ml total of SM buffer (5.8 g NaCl, 2 g $MgSO_4 \cdot 6H_2O$, 50 ml 1 M TrisHCl, pH 7.5, and 5 ml 2% gelatin per liter).

The library was plated as described in Sambrook (supra, pp. 2.61-2.63), on twenty, 140 mm agarose plates at 2.5 x 10^4 plaques per plate. The lambda phage plaques were blotted onto nitrocellulose and treated with denaturing and neutralizing solutions followed by baking at 80°C in a vacuum oven. Filters were then pre-hybridized in 50% formamide, 5 X SSC (75 mM Na citrate; 750 mM NaCl), 0.1% SDS, 5 X Denhardt's

solution (0.1% bovine serum albumin (BSA), 0.1% ficoll, 0.1% polyvinylpyrrolidone), 200 µg/ml yeast tRNA, 100 µg/ml salmon sperm DNA at 42°C for 2 hours. Fragments of human immunoglobulin kappa chain DNA were labeled with a Prime-It® kit (commercially availble from Stratagene) in substantial accordance with the directions provided by the manufacturer, and hybridized with the blots overnight in hybridization solution (50% formamide, 5 X SSC, 0.1% SDS , 1 X Denhardt's solution (0.02% BSA, 0.02% ficoll, 0.002% polyvinylpyrrolidone), 100 µg/ml salmon sperm DNA) at 42°C. The blots were washed twice at 42°C in 2 X SSC and 0.1% SDS for 20 minutes, then at 65°C in 0.2 X SSC, 0.1% SDS for 20 minutes and exposed to XAR-5 X-ray film (commercially available from Eastman Kodak Corp.) overnight at -70°C between two intensifying screens.

The positive plaques were picked and subjected to two rounds of phage DNA purification as described in Sambrook (supra, pp. 2.73-2.76). The purified phage DNA was analyzed by restriction enzyme mapping and Southern blot, as described in Sambrook (supra, pp. 9.31-9.57). **Figure 14** provides a restriction map of the IM9 kappa gene in bacteriophage lambda EMBL3.

b. Subcloning the Intact Kappa Gene into pBluescript®

Southern Blot analysis was used to map the intact kappa chain gene to an 8.8 Kb BamHI fragment. This fragment was isolated from the lambda phage DNA by digestion with BamHI followed by agarose gel electrophoresis. The 8.8 Kb BamHI fragment was ligated using T4 DNA ligase (commercially available from Life Technologies, Inc.) following manufacturers instructions, with pBluescript®SK$^-$ (commercially available from Stratagene, San Diego, CA) which had been previously digested with BamHI. Restriction endonuclease mapping revealed the 5' end of the gene was adjacent to the SacI end of the polylinker.

In order to facilitate modification of the gene, the 5' end of the gene was then sub-cloned as a BamHI to BstEII fragment containing the IgK promoter, the variable exons and a portion of the major intron. The BstEII restriction endonuclease leaves a 5' overhang that is not compatible with any of the sites in the pBluescript®SK⁻ polylinker, so it was necessary to modify the overhanging sequence to make it blunt ended. This was carried out by digesting the pBluescript®SK⁻ clone described above with BstEII and filling in the 5' overhang with Klenow fragment and a solution of all four deoxyribonucleotides, using the method described in Sambrook (supra, pp. 5.40-5.43). This was followed by BamHI digestion and isolation of the 2.2 Kb fragment by agarose gel electrophoresis. This fragment was ligated with pBluescript®SK⁻, previously digested with EcoRV (which leaves a blunt end) and BamHI. The resulting plasmid is shown in **Figure 15**. The DNA sequence of the clone was determined as described in **Example 1**, above.

c. **Engineering the 5' End of the Gene to Create Unique SfiI Sites Flanking the Variable Exon and Removing an MstII Site to Make the Sites Flanking the Constant Exon Unique.**

Two oligonucleotide primers were synthesized on a Millipore DNA synthesizer (Bedford, MA), following manufacturers instructions, for mutagenesis of the 5' end of the IM9 kappa gene: primer B239 (**SEQ. I.D. NO. 31**) TAGTGGATCCAACTGATTTCTCCAT upstream for the BamHI site at the 5' end of the kappa gene and primer B240 (**SEQ. I.D. NO. 32**) TTATTTACTTCTGGGTCACCAGGTTTATTC downstream for the BstEII site in the major intron. The downstream primer recreates the BstEII site that had been altered in the previous step for insertion into pBluescript®SK-.

Two SfiI sites were designed to flank the variable region exon, each having a unique sticky end so as not to re-ligate to each other in cloning but to allow for forced orientation cloning of synthetic variable region cassettes

for CDR grafted antibodies. Each SfiI site, the upstream SfiI site and the downstream SfiI site, involved the design of a pair of oligonucleotide primers and a round of overlap mutagenesis (see **Figure 16**) as described in D.H. Jones and B.H. Howard, "A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction", Biotechniques, 10:62-66 (1991).

Two PCR reactions were performed, each using the variable exon clone as the template. The first used the 5' flanking primer B239 as the 5' primer and the upstream SfiI primer B435 (**SEQ. I.D. NO. 33**) AAGAGGCCGAGCTGGCCCTTCCCTGAATAACCAGGCAGT as the 3' primer. The second used the 3' flanking primer B240 as the 3' primer and the upstream SfiI primer B434 (**SEQ. I.D. NO. 34**) GGGAAGGGCCAGCTCGGCGTGTTCCTATAATATGATCAA as the 5' primer. The products of these reactions were purified and used together as templates in an overlap PCR reaction with primers B239 and B240 as shown in **Figure 16**. The product of the overlap reaction was the full BamHI to BstEII fragment and contained an Sfi site in the appropriate upstream location. This product was used as the template in a new pair of reactions to install the downstream Sfi site in a similar manner, using primers B379 (**SEQ. I.D. NO. 35**) TTCCTGGCCCTGCAGGCCCAGTTGTCTGTGTCTTCTGTT and B380 (**SEQ. I.D. NO. 36**) AACTGGGCCTGCAGGGCCAGGAAGCAAAGTT-TAAATTCTA . The PCR was performed according to the instructions in the GeneAmp® PCR kit (commercially available from Perkin Elmer-Cetus, Norwalk, CT) on a Thermal Cycler® (commercially available from Perkin Elmer Cetus). The reaction was performed for 30 cycles of one minute at 94°C, one minute at 55°C, and two minutes at 72°C in a buffer that contained a 1.5 mM final concentration of $MgCl_2$.

The product of the PCR reaction was cloned into pCR™II vector using a TA Cloning™ Kit (both commercially available from Invitrogen) in substantial accordance with the manufacturer's protocol. The identity of the clone was verified by restriction mapping to be the IM9 kappa BamHI to

BstEII fragment with two engineered SfiI sites of the appropriate size and location.

The MstII site upstream of the kappa promoter, shown in **Figure 16**, was destroyed by linearizing the clone described above with MstII and filling in the 5' overhang to make a blunt end. Re-ligation of the modified ends yielded a sequence that no longer contained an MstII site.

The clone described above was characterized by DNA sequencing analysis as described in **Example 1**.

The engineered BamHI to BstEII fragment was isolated from pCR™II by PCR using two primers, B495 and B496 (**SEQ. I.D. No. 37** CATGTCTGGATCCAACTGATTT and **SEQ. I.D. No. 38** CTGATTTACTTCTGGGTGACCAGGTTTATTCAA respectively).

d. Ligation of the Kappa Gene Fragments with the pSV2gpt (Enhancer minus).

The mutated BamHI to BstEII fragment from the SfiI mutagenesis, described in **Example 8.c.** still contained the native IM9 kappa variable region sequence. It was then ligated with the BstEII to ClaI fragment taken from the pBluescript®SK- clone and the pSV2gpt (enhancer minus) ClaI to BamHI fragment (Beidler, et al, supra).

The resulting clone was analyzed by restriction enzyme mapping, Southern blot analysis, and DNA sequence analysis. The confirmed sequence is provided as a restriction map in **Figure 17**.

e. Insertion of an hZCE Kappa Variable Exon into the pGIM9kappa Vector Using the Engineered SfiI Sites.

The hZCE kappa variable region was taken from a pCR1000™ clone using PCR mutagenesis according to the manufacturer's instructions to add the SfiI sites at the 5' and 3' ends. The oligonucleotide B510 (**SEQ. I.D. NO. 39**)

5'-AAGGGCCAGCTCGGCCT-

CTTCCTATAATATGATCAATAGTATAAATATTTGTGTTTCTATTTCCAATCTCAGGTGCCA

AATGTGACATCCAGATGACCCA-3' was used as the 5' end primer and B511 (SEQ. I.D. NO. 40) 5'-TGGGCCTGCAGGGCCAGGAAGCAAAGTTTAAATTCTAC-TCACGTTTGATTTCCACCTTGGTT-3' as the 3' end primer. The resulting PCR fragment was digested with SfiI. The plasmid pGIM9kappa, deposited with the ATCC with accession number 75512, was also digested with SfiI resulting in three fragments. The hZCE kappa variable region containing fragment described above was ligated with the largest two of the three fragments resulting from the pGIM9kappa digestion. This was carried out as a three fragment ligation reaction. The three SfiI sites have different overhanging sequences due to the nature of the SfiI recognition sequence and so oriented cloning of the three fragments into pGIM9kappa was achieved. The resulting clone pGIM9k/hZCE-kappa was verified by DNA sequence analysis as having the correct Variable exon sequence.

Example 9

Construction and Subcloning of hZCE-CSVL gene.

a. Construction of hZCE-CSVL gene.

The amino acid sequence derived above for the hZCE CDR-grafted CDR switched variable light region was converted into DNA sequence using software from DNA STAR (Madison. WI). Six oligonucleotides with overlapping ends and spanning the sequence of the hZCE-CSV_L gene were synthesized on a Millipore DNA synthesizer (Bedford, MA). The sequences of the six oligonucleotides comprising the template are provided as

SEQ. I.D. Nos. 41-46:

B695 = 5' -GGG-AAG-GGC-CAG-CTC-GGC-CTC-TTC-CTA-TAA-TAT-GAT-CAA-TAG-TAT-AAA-TAT-TTG-TGT-TTC-TAT-TTC-CAA-TCT-CAG-GTG-CCA-AAT-GTG-ACA-TCC-AGA-TGA-CCC-AGT-TTC-CT- 3. (SEQ. I.D. NO. 41)

B696 = 5' -GCA-TGC-CGA-AGT-TGG-AGA-AGG-TGA-AGC-CGG-AGG-CGC-GGC-AGG-TGA-TGT-TCA-CGC-GGT-CGC-CCA-CGG-AGG-CGG-ACA-GGG-TGG-AAG-GAA-ACT-GGG-TCA-TCT-GGA-TGT- 3 **(SEQ. I.D. NO. 42)**'

B549 = 5' -GGC-TTC-ACC-TTC-TCC-AAC-TTC-GGC-ATG-CAC-TGG-ATC-CGC-CAG-AAG-CCC-GGC-AAG-GGC-CTG-AAG-TGG-GTG-GCC-TAC-ATC-TCC-GGC-GGC-TCC-TCC-ACC-GTG-CAC-TA- 3 **(SEQ. I.D. NO. 43)**'

B550 = 5' -GGT-GAT-GGT-CAG-GAA-CAG-CTC-GTT-CTT-GGG-GTT-GTC-GCG-GGA-GAT-GGT-GAA-GCG-GCC-CTT-CAG-GGA-GTC-GGC-GTA-GTG-CAC-GGT-GGA-GGA-GCC-GCC-GGA-GAT-GTA- -3 **(SEQ. I.D. NO. 44)**'

B697 = 5' -CCC-CAA-GAA-CGA-GCT-GTT-CCT-GAC-CAT-CAC-CTC-CCT-GCA-GCC-CGA-CGA-CTT-CGC-CAT-GTA-CTA-CTG-CGC-CCG-CGA-CTA-CTA-CGT-GAA-CAA-CTA-CTG-GTA-CTT-CGA-CGT-GT **(SEQ. I.D. NO. 45)**

B698 = 5' -CAC-AGA-CAA-CTG-GGC-CTG-CAG-GGC-CAG-GAA-GCA-AAG-TTT-AAA-TTC-TAC-TCA-CGT-TTTG-ATC-TCC-ACC-TTG-GTG-CCC-TGG-CCC-CAC-ACG-TCG-AAG-TAC-CAG-TAG-TT **(SEQ. I. D. No. 46)**

The six oligonucleotides were used in a PCR reaction using Taq polymerase and two additional oligonucleotide primers, B553 **(SEQ. I.D. No. 47)** 5' -GGG-AAG-GGC-CAG-CTC-GGC-CTC-TT -3' and B554 **(SEQ. I.D. No. 48)** 5'-CAC-AGA-CAA-CTG-GGC-CTG-CA- 3' for amplification. The oligonucleotide templates, primers, PCR reagents and buffers were used at concentrations described by the manufacturer. Twenty five cycles of amplification were carried out, as follows: (1) Denature at 94°C for one minute, anneal at 55°C for one minute, and extend at 72°C for one minute.

b. Subcloning of hZCE-CSVL gene into TA Vector.

Following PCR synthesis of the CDR-grafted variable region containing ZCE-025 heavy chain CDRs, the approximately 500 base pair DNA fragment was ligated into a TA holding vector as per the manufacturer's protocol (In Vitrogen, San Diego). TA vectors are provided by the manufacturer as

linear molecules containing a single deoxythymidylate as an overhang on each of the vector's 3' ends. This is complementary to the deoxyadenylate overhangs found on the 3' ends of PCR products due to the terminal transferase activity of Taq polymerase.

TA clones containing inserts of the correct size (about 500 base pairs) were identified by EcoRI restriction digests of DNA minipreps using methods known in the art. Up to ten clones with appropriate insert sizes were sequenced on a Genesis® DNA sequencer (DuPont, Delaware, MD). A clone with the appropriate sequence was digested to completion with SfiI restriction endonuclease. This restriction site was present at the 5' and 3' ends of the hZCE-CSV_L gene for cloning into the final expression vector as described in **Example 10**, below. The hZCE-CSV_L fragment was isolated following electrophoresis using the gel purification method described above. After ethanol precipitation, the fragment was resuspended in sterile distilled H_2O and the concentration was determined by running a small aliquot on a gel, as described previously.

Example 10

Construction of Expression vector - pGIM9k/hZCE(CSV_L)-kappa and Expression of hZCE(CSVL)-kappabody

a. Construction of pGIM9k/hZCE(CSV_L)-kappa.

The 484 bp DNA SfiI to SfiI fragment containing the hZCE-CSV_L region was combined with a 9 kb SfiI to SfiI fragment isolated from the pGIM9 kappa expression vector by standard ligation (Sambrook, *et al.*). As shown in **Figure 18**, the resulting expression vector, pGIM9k/hZCE(CSV_L)-kappa contained the following components:

(1) Human IM-9 kappa promoter, signal exon 1 and signal intron (up to added SfiI site).

(2) The hZCE(CSV_L) gene beginning with an SfiI

site in the signal intron and including the pGIM9 kappa signal exon II hZCE(CSV_L) region and extending to an SfiI site at beginning of the major intron.

(3) Human IM-9 kappa major intron (from SfiI site), kappa constant exon and 3' flanking sequences (containing native polyadenylation site).

(4) XGPRT gene under the control of an enhancerless SV40 early promoter.

(5) Bacterial plasmid origin of replication, derived from pBR 322.

(6) Bacterial ß-lactamase, driven off its native promoter.

b. Transfection of SP 2/0 and hZCEK.

Vector pGIM9k/hZCE(CSV_L)-kappa, on deposit with ATCC under the provisions of the Budapest Treaty Deposit No. 75530, was electroporated into two different host cell lines, SP 2/0 and hZCEk. hZCEk is a transfectoma derived from SP 2/0 by transfection with the vector pGIM9k/hZCE-kappa, which expresses CDR grafted ZCE/IM-9 light chain (hZCEK-homodimer) [**Example 8.e.**]. For SP 2/0, pGIM9k/hZCE(CSV_L)-kappa was electroporated together with the drug selectable gene neo in the vector pSV2Neo, and transfectants were selected by growth in HH4 medium containing 1.5 mg/ml geneticin (Bethesda Research Labs/Gibco, Gaithersberg, MD). For hZCEk, pGIM9k/hZCE(CSV_L)-kappa was also co-electroporated with pSV2neo to allow selection of transfectants in medium with geneticin 1.5 mg/ml. Electroporation conditions and selection media recipes were as described by Chu, *et al.* (Nucleic Acids Research, 15:1311-1325 (1987)). Briefly, the SP2/0 cells were grown in media containing 10% FBS and were maintained in log phase growth for the three days preceding electroporation. Fifty micrograms of the plasmid vector was linearized using the restriction enzyme PvuI (1 unit/µg) and the Reaction Buffer

#7 from GIBCO-BRL (Gaithersburg, MD). At the time of transfection the SP2/0 cells were collected by centrifugation in an IEC clinical centrifuge (800 rpm, 10 min, room temperature). Cells were washed in Hanks Buffered Saline Solution from Gibco Laboratories (Grand Island, NY) containing an additional 6 mM dextrose and resuspended at a final concentration of 1.0×10^7 cells/ml. 0.5 ml of cells were aliquoted into cuvettes and the linearized DNA was added. Electroporation was done using the Cell-Porator® (GIBCO-BRL) with settings of 300 µF and 350 volts.

c. Selection and characterization of hZCE-kb expressing clones.

Resistant clones of each host cell line were identified by growth on appropriate selective media and assayed for hZCE(CSV_L) chain production (SP 2/0 host) and CEA binding (hZCEk host) activity as described in **Example 15**, shown below. The resultant clones were called hZCEhb (SP 2/0 host) and hZCEkb (hZCEk host). hZCEhb produces only the human kappa light chain with ZCE heavy chain CDRs secreted as a homodimer, while hZCEkb produces a human light chain dimer with one kappa chain containing ZCE heavy chain CDR's and the other containing ZCE light chain CDRs. A conventional human kappa ELISA can be used to quantitate production levels of the homodimer from hZCEhb, but a CEA-binding ELISA is required to quantitate the antigen binding heterodimer hZCEkb. The hZCEkb chain or hZCEhb chain were secreted as dimers. The hZCEhb homodimer did not bind CEA, while the hZCEkb had affinity for CEA.

Example 11

Construction of hZCE(CSV$_L$) Expression Vector and Expression of hZCE(CSV$_L$)

a. Preparation of gene for chelating peptide.

A CDR switched variable region isolate was constructed as a variation of the hZCE(CSV$_L$) kappa chain where the human kappa constant region would be deleted so as to express the hZCE(CSV$_L$) light chain domain only. To screen for the CDR switched isolate construct, it was desirable to express it as a fusion protein containing a metal chelating peptide for purification. The gene encoding the chelating peptide was prepared by creating a DNA fragment which would ultimately replace the human kappa constant exon in the pGIM9k/hZCE(CSV$_L$)-kappa vector. Using PCR techniques, an approximately 330 base pair MstII/MstII modified fragment (Fragment A) was prepared using the pGIM9k/hZCE(CSV$_L$)-kappa expression vector as template. The upstream primer in this PCR reaction was B1000 (**SEQ. I.D. No. 49**) 5'-CAC-CAT CCT GTT TGC TTC TTT CCT CAG GAA CTG TGC ACT GGC ACC ACC ACC CAT AGA GGG AGA AGT GCC CCC ACC TGC TCC TCA GTT -3', which included the codons for a 6-amino acid chelating peptide, and the downstream primer was B441 (**SEQ. I.D. No. 50**) 5'-GGGTAAAAATAGAATGAAGGATGAT-TTTTATAAAT-3'. Fragment A consisted 5' to 3' of (1) an MSTII restriction site and the splice acceptor site from the IM9 kappa constant region; (2) the codons for the first three amino acids of the kappa constant region; (3) the codons for a six amino acid chelating peptide sequence (HWHHHP) and a termination codon; and (4) 3' untranslated sequence including the polyadenylation site and native MSTII-restriction site.

b. Construction of pGIM9k/hZCE(CSV$_L$) expression vector.

Fragment A and pGIM9k/hZCE(CSV$_L$)-kappa were digested with either MstII or Bsu36-1 (Stratagene, 10X

Universal buffer, 37°C for a minimum of 3 hours) to produce ligatable ends. Fragments (~330 bp of Fragment A and ~12.8 kb pGIM9k/hZCE(CSV_L)-kappa were thus isolated and purified using Milligen's Ultrafree-MC (Yonezawa, Japan) method. Ligation was carried out using components and ligation conditions from a TA Cloning Kit (Invitrogen, San Diego, CA) following the manufacturer's protocol. Electroporation into Electromax DH10B cells (BRL, Gaithersburg, MD) was performed. Transformed cells were plated onto agar, incubated overnight, and colonies were grown-up for plasmid mini preps using Qiagen's (Chatsworth, CA) "Mini Plasmid" protocol. Construct size was verified by restriction digest analysis using EcoRI, MstII or Bsu36-1, SstI, and BamHI enzymes.

Large scale plasmid preparations were performed using Qiagen's "Maxi-plasmid" prep procedure. DNA sequencing was performed to verify the correct sequence, which is called hZCE(CSV_L) (**SEQ I.D. No. 51**). The cloned plasmid herein is called pGIM9k/hZCE(CSV_L).

SEQ. I.D. No. 51

GAC ATC CAG ATG ACC CAG TTT CCT TCC ACC CTG TCC GCC TCC GTG
GGC GAC CGC GTG AAC ATC ACC TGC CGC GCC TCC GGC TCC ACC TTC
TCC AAC TTC GGC ATG CAC TGG ATC CGC CAG AAG CCC GGC AAG GGC
CTG AAG TGG GTG GCC TAC ATC TCC GGC GGC TCC TCC ACC GTG CAC
TAC GCC AAC TCC CTG AAG GGC CGC TTC ACC ATC TCC CGC GAC AAC
CCC AAG AAC GAG CTG TTC CTG ACC ATC ACC TCC CTG CAG CCC GAC
GAC TTC GCC ATG TAC TAC TGC GCC CGC GAC TAC TAC GTG AAC AAC
TAC TGG TAC TTC GAC GTG TGG GGC CAA GGG ACC AAG GTG GAA ATC
AAA

c. Expression of hZCE(CSV_L).

Linearization of pGIM9k/hZCE(CSV_L) DNA was performed via ClaI digestion. Electroporation into SP2/0 cells was performed as previously described in **Example 10.b.** Cells were seeded in HH4 medium supplemented with 10% FCS. Three days later, cells were plated @ 2 x 10 @ 5/ml in 24-

well cluster plates in the presence of HH4, 10% FCS, MAX (MAX = 1.0 µs/ml mycophenolic acid plus 100 µs/ml xanthine). At day 14 after plating, colonies were harvested and transferred to 6-well plates (Falcon) for expansion and serum-free medium adaptation. Clones were successfully expanded and adapted to serum-free conditions within 2 weeks.

Example 12

Construction of pNIM9k/hZCE-gamma

a. Construction of hZCE heavy chain variable exon.

The protein sequence of the heavy chain of hZCE was converted to nucleic acid sequence in the following manner: (1) if the amino acid was derived from ZCE, the actual ZCE codon at the site was used; (2) if the amino acid was derived from IM9, the actual IM9 codon at the site was used; (3) if the amino acid was derived from a consensus sequence, any appropriate codon was used.

The hZCE gamma variable exon (**SEQ. I.D. NO. 58**) shown below was obtained by PCR reactions.

SEQ I.D. No. 58

GAA ATG CAA CTG GTG GAA TCT GGG GGA GGC CTG CTA CAG CCT GGC
CGG GCC CTG CGG CTC TCC TGT GCA GCC TCT GGA TTC ACT TTT AGT
AAC TTT GGA ATG CAC TGG ATT CGG CAA ACT CCA GGG AAG GGC CTG
GAG TGG GTC GCA TAC ATT AGT GGT GGC AGT AGT ACC GTC CAC TAT
GCA GAC TCC TTG AAG GGC CGA TTC ACC ATC TCC CGG GAC AAC GCC
AAG AAC TCC CTC TAT TTG CAA ATG ACC AGT CTC CGG GCT GAG GAC
ACG GCC TTG TAT TAC TGT GCA CGG GAT TAC TAC GTT AAT AAC TAC
TGG TAC TTC GAT GTC TGG GGC CAA GGG ACA ATG GTC ATC GTC TCT
TCA G

Five overlapping oligonucleotides, B156, B159, B396, B397, and B398 (**SEQ. I.D. NO. 60-64**)

SEQ I.D. No. 60

5'-GAT CCG AAA TGC AAC TGG TGG AAT CTG GGG GAG GCC TGC TAC AGC CTG GCC GGG CCC TGC GGC TCT CCT GTG CAG CCT CTG GAT TCA CCT TTA G-3'

SEQ I.D. No. 61

5'-CAC CAC TAA TGT ATG CGA CCC ACT CCA GGC CCT TCC CTG GAG TTT GCC GAA TCC AGT GCA TTC CAA AGT TAC TAA AGG TGA ATC CAG AGG C-3'

SEQ I.D. No. 62

5'-GGG TCG CAT ACA TTA GTG GTG GCA GTA GTA CCG TCC ACT ATG CAG ACT CCT TGA AGG GCC GAT TCA CCA TCT CCC GGG ACA ACG CCA AGA A 3'

SEQ I.D. No. 63

5'-TAT TAC TGT GCA CGG GAT TAC TAC GTT AAT AAC TAC TGG TAC TTC GAT GTC TGG GGC CCA GGG ACA ATG GTC ATC GTC TCT TCA -3'

SEQ I.D. No. 64

were synthesized on a DNA Synthesizer (Millipore) following manufacturer's instructions. They were fused together by a PCR reaction using B161 (**SEQ. I.D. NO. 65**) 5'-AAG GAT CCG AAA TGC AAC TGG TGG AAT CT -3' and B162 (**SEQ. I.D. NO. 66**) GAC GAA TTC TGA AGA GAC GAT GAC CAT TG as the end primers. The resulting fused fragment was cloned into pCR™II (Invitrogen) and the sequence was verified as described in Step 2.1.j.

b. Construction of the hZCE gamma expression vector, pNIM9k/hZCE-gamma (cDNA).

The hZCE heavy variable exon and the entire IM9 gamma constant region (from 5' IM9 heavy CH1 exon to the BstEII site 3' of the CH3 exon) were fused together by an

overlap PCR reaction. Two PCR reactions were performed: the first PCR reaction used the pCRII clone from 3.a. as template and primers B611 and B612. The PCR product was reamplified with primers B467 and B567. The second PCR reaction used primers B566 and B514. The products of these two reactions were used together as templates in the overlap reaction with primers B467 and B514. The resulting fusion fragment of hZCE heavy variable exon and 5' IM9 heavy CH1 exon to BstEII was cloned into pCR™II (Invitrogen) and the sequence was verified as described in Step 3.1. The resulting vector is phZCE/CHIBstEII.

A pair of oligonucleotides, B743 and B744, were designed to add the splice recognition site and the SfiI site 3' of the variable region. The IM9 heavy chain cDNA vector was digested with BamHI and HindIII, extracted with phenol and chloroform mixture, precipitated with EtOH, and resuspended in TE. Primers B743 and B744 were kinsed, annealed together, and ligated with the digested vector. The ligation reaction was used to transform E. coli DH10B by electroporation. The colonies were picked for analysis by restriction enzyme mapping and the resulting vector is pIM9gammacDNASfiI.

The phZCE/CH1BstEII vector and pIM9gammacDNASfiI were digested with SfiI and BstEII. The 740 bp fragment from phZCE/CH1BstEII and the 950 bp fragment were purified by agarose gel electrophoresis. The pGIM9kappa vector was digested with SfiI and the 12 Kb fragment was purified by agarose gel electrophoresis. The three purified fragments were ligated and used to transform E. coli DH10B by electroporation. The colonies were picked for analysis by restriction enzyme mapping. The resulting vector is pGIM9k/hZCE-gamma .

The Neomycin resistance gene was inserted into pGIM9kappa vector to make pNIM9kappa. Both the pGIM9kF2 and the pSV2neo vectors were digested by ApaI and PvuI, the 5 Kb Neomycin resistance gene-containing fragment from the pSV2neo digest and the 9 Kb fragment from the pGIM9k digest were

purified by gel electrophoresis. The two fragments were ligated and used to transform E. coli DH10B by electroporation. The colonies were analyzed by restriction enzyme mapping, the resulting plasmid is pNIM9kappa.

Both the pNIM9kappa and the pG(IM9k)/hZCE-gamma vectors were digested with SfiI, the 9 Kb and 5 Kb fragments from pNIM9kappa and the 1.6 Kb fragment from pG(IM9k)/hZCEgamma were purified by agarose gel electrophoresis. The three purified fragments were ligated and used to transform E. coli DH10B. The colonies were picked and analyzed by restriction enzyme mapping, the resulting plasmid is pN(IM9k)/hZCE-gamma(cDNA).

Example 13

Construction of hZCE(CSV_L)-gamma Expression Vector and expression of hZCE(CSV_L)-intact kappabody

In another variation using the variable kappa region from IM-9 containing ZCE heavy chain CDRs (hZCE(CSV_L) region), a human gamma heavy chain was constructed.

a. Construction of pGIM9k/hZCE(CSV_L)-gamma

Using the polymerase chain reaction (PCR), a 2.1 kilobase DNA fragment was amplified using primers B922 (**SEQ. I.D. No. 52**) 5'-AAG-AGC-TCC-TGA-ACC-TCG-CGG-ACA-GTT-AA-3') and B923 (5'-AAA-TCG-ATC-TCA-GGC-CTC-AGA-CTC-GGC-CTG-ACC-CGT-GGA-AA-3') (**SEQ. I.D. No. 53**) from a fragment of pNIM9k/hZCE-$gamma_1$. The 5' end of this fragment contained an Sst-1 restriction site and the 3' end contained a Cla-1 site. A second Cla-1 (5') to Sst-1 (3') fragment of 8.5 kilobases containing the neomycin gene, ß lactamase gene and the hZCE-CSV_L variable region gene was ligated together with the PCR generated 2.1 kilobase fragment. The 10.6 kilobase plasmid resulting from this ligation was reopened with Sst-1 restriction endonuclease and ligated together with a 2.2 kilobase Sst-1 fragment from pGIM9kappa containing a portion

of the human kappa major intron with enhancer. The final expression vector is 12.8 kilobases and called pNIM9k/hZCE(CSV_L)-gamma.

b. Expression of hZCE(CSV_L)-intact kappabody.

The pNIM9k/hZCE(CSV_L)-gamma expression vector was electroporated (as described in **Example 10.b.**) into cells expressing the pGIM9k/hZCE-kappa gene. Three days following electroporation the cells were put under drug selection (geneticin 1.5 mg/ml) and colonies which grew up under this selection were analyzed for secretion of hZCE(CSV_L)-intact antibody.

Protocol for Subcloning Transfectomas

Individual wells from the initial screening for cells secreting the highest levels of immunoglobulin were further subcloned to insure a single clone had been selected. Briefly, the cells were diluted to 10, 5 or 0.3 cells per 200µl and plated into two 96-well tissue culture plates at each dilution. The medium is HH4 with 10% fetal calf serum, 100 ug/ml xanthine and the appropriate selection drug. After fourteen days individual wells were visually screened for single colonies, then harvested and cultured further so as to obtain a quantitative ELISA value as described in **Example 15**, below.

Example 14

Cloning and Expression of a Single Chain Fv Containing a CSV_L Fragment

a. Construction of pGIM9k/hZCE(CSV_L)-ScFv expression vector.

To construct an expression vector for a CSV_L, the earlier expression vector pGIM9k/hZCE-kappa was reconstructed to contain the cdr-grafted kappa variable region in place of the human kappa constant region. In addition, the vector contained a 5' extension to the kappa variable region to

serve as a linker (L) between variable regions as well as a 3' chelating peptide (CP) sequence. Therefore, diagramatically, the linear construct is as follows:

The kappa variable region with 5' linker and 3' chelating peptide was synthesized in three separate PCR reactions. The first DNA fragment (Fragment-1) of 457 base pairs was amplified from vector pGIM9k/hZCE-kappa (**Example 10.a.**) with a 5' primer (C101) (**SEQ I.D. No. 54**) 5'-CTG TTT GCT TCT TTC CTC AGG AGG CGG TTC AGG AGG ATC AGG CGG TTC AGG TGG ATC AGG AGG CGA CAT CCA GAT GAC CCA GTC TCC T-3' containing the MstII restriction site, the linker $(G\text{-}G\text{-}S)_4GG$ and the first 24 bases of the constant kappa gene (**Example 10.a.**); and a 3' primer C102 [**SEQ I.D. No. 55**] 5'-GTC AGG CTG GAA CTG AGG AGC AGG TGG GGG CAC TTC TCC CTC TAT GGG TGA TGG TGC CAA TGT TTG ATT TCC ACC TTG GTC CCT TGG CCG -AA-3' containing the bases of the 3' end of the kappa constant region, a H-W-H-H-H-P chelating peptide and stop codon. The second DNA fragment (Fragment-2) of 335 base pairs was generated using 5' primer C103 [**SEQ I.D. No. 56**] 5'-GAG AAG TGC CCC CAC CTG CTC CTC AGT TCC AGC CTG ACC CCC TCC CAT CCT -3' and 3' primer B441 [**SEQ I.D. No. 50**] and the same template as above, i.e. pGGhZCE-HB. This Fragment contained the 3' human kappa constant region containing the polyadenylation signal.

The final DNA fragment (Fragment-3) was amplified using Fragment-1 and Fragment-2 as template and 5' primer C101 [**SEQ I.D. No. 54**] and 3' primer B441 [**SEQ I.D. No. 50**] to give the approximately 800 base pair Fragment-3. This Fragment-3 was cloned into a TA vector for confirmation of DNA sequence as described in **Example 9.b.** Following confirmation of sequence the Fragment-3 insert was re-isolated from the TA vector as an MstII fragment and cloned into the vector pGIM9k/hZCE-hb (which had its MstII fragment, containing the human kappa constant region, deleted). All PCR amplifications were carried out as described in **Example 9.a.**

b. Expression of hZCE(CSV_L)-ScFv.

After cloning and scale up, the final expression vector, herein called pGhZCE-CSV_L-sFV, was electroporated into SP2/0 hybridoma cells as described in **Example 10.b.** Clones secreting the CSV_L-sFV construct were identified as described in **Example 15.f.**, below. Finally, the affinity of the construct was analyzed via a competitive inhibition assay as described in **Example 15.e**, below.

Example 15

Identification, quantitation and affinity determinationof engineered constructs produced

a. Identification and quantitation of secreted hZCE-CSVL-kappa homodimer and hZE(CSVL)/hZCE-kappa heterodimers

Identification and quantitation of secreted CDR grafted human kappa chains from transfected SP 2/0 cells expressing hZCE kappa homodimer, and those expressing hZCE-CSV_L homodimer were identified by a standard enzyme-linked immunosorbent assay ("ELISA", as described by Engvall, E. and Perlmann, P., Immunochemistry, 8:871-874 (1971)) for human kappa. The purpose of this assay was to identify those cells secreting the highest levels of kappa chain polypeptide coded for by pGIM9k/hZCE-kappa or pGIM9k/hZCE(CSV_L)-kappa plasmid vector. A 5µg/ml solution of goat anti-human kappa chain (Tago #4106, Tago Inc., Burlingame, CA) in 10mM sodium phosphate pH 7.4 was prepared. Each well of a 96 well plate was coated with 50µl of this solution. The plates were then incubated overnight at 37°C. Plates were then rinsed thoroughly in H_2O, and then PBS with 1.0% Tween-20™ (w/v). Fifty µl of the supernatant fractions were added to each well, and incubated for two hours at room temperature. Plates were again rinsed as detailed above. A goat anti-human kappa chain alkaline phosphatase conjugate (Tago #2496,

Tago, Inc.) was diluted 1:1000 in the same medium as the supernatant material. 100µl were added per well and allowed to incubate for one hour at room temperature. Plates were rinsed as above. The alkaline phosphatase substrate (Hybritech, Inc., San Diego, CA Part #100103) was prepared as per package instruction, one tablet per 3ml of distilled H_2O and 150µl of this substrate was added to each well and allowed to incubate 30 minutes at 37°C. The reaction was quenched with 50µl of 300 mM EDTA and then the absorbance was read at 405 nM. Colonies, whose supernatants showed the highest levels of kappa expression, were subcloned and cryo-preserved. Expression levels are shown in **Table 3**.

b. Identification and quantitation of hZCE(CSV_L)-intact kappabodies.

Detection of assembled hZCE(CSV_L)-intact kappabodies was carried out by coating the microtiter plate wells with goat anti-human IgG heavy chain antibody reagent (Tago #3100, Tago, Inc., 887 Mitten Road, Burlingame, CA) at 5 µg/ml in 10 mM phosphate pH 7 to 8. Plates were dried overnight at 37°C, then washed with PBS and 0.1% Tween-20™, then H_2O. Fifty microliters of the cell supernatant were added to each well and incubated for 2 hours at room temperature. Plates were again rinsed as detailed above. A goat anti-human kappa chain alkaline phosphatase conjugate (Tago #2496 Tago, Inc., 887 Mitten Road, Burlingame, CA) was diluted 1:1000 in the same medium as the supernatant material. 100 µl were added per well and allowed to incubate for 1 hour at room temperature. Plates were rinsed as above. The alkaline phosphatase substrate (Hybritech), one tablet per 3 ml of distilled H_2O, and 150 µl of this substrate was added to each well and allowed to incubate 30 minutes at 37°C. Purified protein, IgG_1-kappa, from the human lymphoblastoid cell line IM9 was used as a positive control.

c. ELISA for detecting the presence of hZCE(CSV_L) constructs bound to Carcino-Embryonic Antigen (CEA).

To detect the hZCE(CSV_L) constructs which can bind CEA, ELISAs were performed as follows:

On the first day CEA stock standard (1mg/ml) (Hybritech Part # 211288) was diluted to 10 µg/ml in PBS with 1mg/ml BSA in a final volume of 6 mL for each ELISA plate. 96-well ELISA plates (Titertek, McLean, VA) were coated at 50µl/well, tapped to ensure that all well bottoms were completely covered, and incubated overnight at 37°C.

On the second day the plates were washed twice with distilled, deionized water, twice with 1XPBS+0.1% Tween-20™, and twice again with distilled, deionized water. Samples containing the hZCE(CSV_L)-heterodimer, hZCE(CSV_L)-intact, and standards were added to the plates at 50µl/well. Plates were then sealed and incubated at room temperature for 2 hours. Goat Anti-(Human Kappa) conjugated with alkaline phosphatase (Tago # 2496 Burlingame, CA) was diluted 1:1000 in RPMI medium (Gibco) with 10% horse serum and 3% goat serum to a volume of 10ml/plate. Plates were washed as before, and the anti-Kappa conjugate was added to the plates at 50µl/well. Then the plates were sealed and incubated at room temperature on a shaker at ~100 rpm. for 1 hour.

PNPP Alkaline Phosphatase tablets (Hybritech Part #100103) were dissolved in distilled, deionized water at a ratio of 1 tablet per 3ml of water, and 150µl/well of the alkaline phosphatase solution was added to the plates. The plates were incubated at 37°C for half an hour and then absorbencies were read at 405nm using a CERES900 ELISA reader (BioTek, Inc., Winoosk, Vermont). The cultures corresponding to the wells whose supernatants yielded the highest optical densities were selected for further scale up and ELISA quantitation.

d. Competitive inhibition ELISA for quantifying CEA binding of hZCE CSV_L heterodimers.

The binding affinity of the hZCE(CSV_L)-heterodimers for carcinoembryonic antigen was quantified as follows:

On the first day the substrate antibody was prepared. Briefly, CEV124.1, a murine monoclonal anti-CEA antibody obtained from Hybritech (San Diego, CA) was diluted 1:1000 in phosphate buffered saline (PBS) to a final volume of 6 mL. The PBS was prepared by mixing 1494 g NaCl, 36 g KCl, 36 g KH_2PO_4, and QS to 18L H_2O, then diluted 1:10 with distilled, deionized water. A 96 well plate was coated with the antibody-containing solution using about 50µl/well. The plate was tapped to ensure that each entire wall bottom was covered. The plate was sealed and left at room temperature overnight. The next day the CEA antigen was prepared as described in **Example 15.C.**

The plates containing bound antibody were washed four times with distilled, deionized water, and 50µl of the CEA/BSA antigen-containing solution was dispensed into each well. The plates were sealed and placed on a rotator shaking at ~300 rpm. for 2 hr. Finally the plates were washed as before.

A supernatant of hZCE(CSV_L)-heterodimer was loaded at 50µl/well. A standard curve was generated by diluting a 10µg/ml solution of XCEM F(ab)' or ZCE Fab' at 1:2 increments along the top row of the assay plate. The XCEM chimeric antibody was described in Beidler, C.B., *et al.*, "Cloning and High Level Expression of a Chimeric Antibody with Specificity for Human Carcinoembryonic Antigen," J. of Immunol., 141:4053-4060 (1988). Plates were sealed and incubated on a rotator as before for 45 minutes to allow the test antibody to bind to the antigen.

For use as the competition antibody, biotinylated F(ab)' fragments of XCEM chimeric monoclonal antibody or ZCE Fab' were prepared. Biotinylation was conducted as described by Enzotin Biochem, Inc., New York, NY. The biotinylated fragments were diluted to a final concentration of 0.4µg/ml

(experimentally determined to give an OD 490 of about 0.6) and 50µL of the solution was added to each well without washing the plate. The plate was sealed and incubated as before for 45 minutes and then washed as before.

Streptavidin/horse radish peroxidase (Fisher Biotech, Pittsburgh, PA) conjugate was prepared as per manufacturer's directions and then diluted in 1X PBS with 1% BSA. Fifty µl of the streptavidin labeled conjugate were added to each well, the plate was sealed, and incubated as before for 45 minutes. Finally the plate was washed as before to remove the unbound conjugate.

Final substrate was made by completely dissolving one 10mg tablet of o-phenylenediamine dihydrochloride (Sigma #P8287, St. Louis, MO.) into 10 ml of PCB (18.45g Citric Acid (monohydrate), 25.86g Na_2HPO_4, bring to 1.81 @ MilliQ H_2O, pH to 5.0, QS to 2L), then adding 15µl of 30%H_2O_2. 100µl of substrate was added to each well. When a standard curve could be visualized, the plates were quenched by adding 50µl of 4M H_2SO_4 to each well. The assay was read on the Biotek CERES900 assay reader at absorbance of 490 nM. Concentration calculations were done using the built-in software "Kineticalc Jr." from BioTek Instruments (Winoosk, Vermont). Results of these experiments are shown in **Table 3** below.

e. Competitive inhibition assay for determination of affinity of anti-CEA antibodies and constructs.

Affinities of unlabeled recombinant antibodies were determined by a modification of the method described by H. Motulsky and L. Mahan, *Molecular Pharmacology*, 25:1-9, 1983). This method can measure the affinity of unlabeled antibodies by evaluating their ability to inhibit the binding of a labeled tracer antibody which reacts with the same epitope of an antigen.

Tandem® R CEA Beads (Hybritech #600211), which contain the mouse antiCEA antibody CEV124, were put into 13cm x 75cm polystyrene tubes (1 bead per tube) and incubated with

100mg of CEA, diluted in 1% BSA/PBS solution to a final volume of 100ul, for 2-5 hours at room temperature. The source of the CEA used for these experiments is CEA Stock Standard Solution (Hybritech, #200288). The beads were then washed twice with 2ml of 0.1% Tween20™ in phosphate buffered saline just prior to adding the antibodies for affinity testing.

The tracer antibody is a isothiobenzyl-DTPA conjugate of ZCE025 Fab' fragment labeled with 3uCi of ^{111}In Citrate per microgram of Fab'. The tracer is first titrated for binding to the above CEA beads to determine a 40-60% saturation point. This concentration of tracer (usually 1.5 x 10^{-9} M) is used for all the following inhibition reactions. Varying concentrations of unlabeled XCEM or supernatant containing hZCE-CSV_L heterodimer were added (100ul) to the CEA beads at 2X their final concentrations (final is 1 x 10^{-7} M down to 1 X 10^{-11} M, diluted in 1%BSA/PBS) together with an equal volume of the 2X tracer (100 µl). The reaction was then incubated overnight at room temperature on an Orbital Shaker (150-200 RPM).

f. Identification and quantification of hZCE(CSV_L) isolate.

Cells putatively secreting hZCE(CSV_L)-isolate were seeded at 4 x 10^5/ml in serum-free HH4 medium containing 100 µg/ml xanthine and 1.0 µg/ml mycophenolic acid. When cell numbers reached ~1 x 10^6/ml, 1.0 ml of their supernatants were collected and mixed with 100 µl of Ni^{+2}-loaded nitrilo acetic acid agarose beads (Qiagen, Inc., Chatsworth, CA). The beads and conditioned cell supernatant from 24 individual clones were incubated for a minimum of four hours on a rotating wheel at room temperature. The beads were washed 3 times with 50 mM sodium phosphate, 100mM sodium chloride buffer, pH 7.4. Bound protein was eluted from the beads by addition of 100 µl of SDS-PAGE reduced sample buffer. The elutate was electrophoresed on 15 - 20% SDS-PAGE gels and the gels were silver stained to visualize and quantitate the

hZCE-CSV_L-isolate. The SDS-PAGE gels, buffers and silver staining kit were carried out using reagents from Biorad, (Richmond, CA) according to the manufacturer's instructions. Results are shown in **Table 3** below.

TABLE 3

Construct	Secreted Protein µg/ml*	Affinity 1/M	M.W. (Kd)
mZCE Fab	EP	2×10^9	52
hZCE-kappa homodimer	20	0	50
hZCE(CSV_L)-kappa homodimer	20	0	54
hZCE(CSV_L)-kappa heterodimer	20	2×10^9	52
hZCE(CSV_L)-intact	3	ND	160
hZCE(CSV_L)	1	ND	18

* = average
EP = enzymatically produced
ND = not determined

The foregoing description of the invention is exemplary for purposes of illustration and explanation. It should be understood that various modifications can be made without departing from the spirit and scope of the invention. Accordingly, the following claims are intended to be interpreted to embrace all such modifications.

GGSGGSGGSGGSGG (14 RESIDUES) (**SEQ I.D. No. 1**)

HWHHHP (6 RESIDUES) (**SEQ I.D. No. 2**)

5'GAC TAG CGG CCG CAT CGA TCC CCC CCC CCC CCC C (**SEQ. I.D. No. 3**)

5'CAG ACG TCG ACG ATG GAT ACA GTT GGT GCA GCA TC (**SEQ. I.D. No. 4**)

SEQ. I.D. NO. 5

ZCE-025 Light Chain Variable cDNA

GAC ATT GTG ATG ACC CAG TCT CAA AAA TTT ATG TCC ACA TCA GTT
GGA GAC AGG GTC AAC ATC ACC TGC AAG GCC AGT CAG AAT GTT CGT
ACT GCT GTA GCC TGG TAT CAA CAG AAA CCA GGG CAG TCT CCT AAA
GCA CTG ATT TAC TTG GCA TCC AAC CGG TAC ACT GGA GTC CCT GAT
CGC TTC ACA GGC ATT GGA TCT GGG ACA GAT TTC ACG CTC ATC ATT
AGC AAT GTG CAA TCT GAA GAC CTG GCA GAT TAT TTC TGT CTG CAA
CAT TGG AAT TAT CCT CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG
CTG AAA C
381

SEQ. I.D. No. 6

DIVMTQSQKFMSTSVGDRVNITCKASQNVRTAVAWYQQKPGQSPKALIYLASNRYTGVPDR
FTGIGSGTDFTLIISNVQSEDLADYFCLQHWNYPLTFGAGTKLELK

5'CAG ACG TCG ACG TTC CAG GTC ACT GTC ACT GGC TC (**SEQ. I.D. NO. 7**)

SEQ. I.D. NO. 8

ZCE-025 Heavy Chain Variable cDNA Sequence:

GAT GTG CAG CTG GTG GAG TCT GGG GGA GGC TTA GTG CCG CCT GGA
GGG TCC CGG AAA CTC TCC TGT GCA GCC TCT GGA TTC ACT TTC AGT
AAC TTT GGA ATG CAC TGG ATT CGT CAG GCT CCA GAG AAG GGA CTG
GAG TGG GTC GCA TAC ATT AGT GGT GGC AGT AGT ACC GTC CAC TAT
GCA GAC TCC TTG AAG GGC CGA TTC ACC ATC TCC AGA GAC AAT CCC
AAG AAC ACC CTG TTC CTA CAA ATG ACC AGT CTA AGG TCT GAA GAC
ACG GCC ATG TAT TAC TGT GCA AGA GAT TAC TAC GTT AAT AAC TAC
TGG TAC TTC GAT GTC TGG GGC GCA GGG ACC ACG GTC ACC GTC TCC
TCA G
420

SEQ. I.D. NO. 9

DVQLVESGGGLVPPGGSRKLSCAASGFTFSNFGMHWIRQAPEKGLEWVAYISGGSSTVHYA
DSLKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARDYYVNNYWYFDVWGAGTTVTVSS

SEQ. ID NO 10

GAC ATC CAG ATG ACC CAG TTT CCT TCC ACC CTG TCT GCT TCT GTA
GGA GAC AGA GTC ACC 60
ATC ACT TGT CGG GCC AGT CAG AGT ATT AGT GCC TGG TTG GCC TGG
TAT CAG CAG AAA CCA 120
GGG AAA GCC CCT AAA CTC CTG ATC TAT AAG GCG TCT AGT TTA GAA
AGT GGG GTC CCA TCA 180
AGG TTC AGC GGC AGT GGA TCT GGG ACA GAG TTC ACT CTC ACC ATC
ACC AGC CTG CAG CCT 240
GAT GAT TTT GCA ACT TAT TTC TGC CAA CAC TAT AAT CGA CCG TGG
ACG TTC GGC CAA GGG 300
ACC AAG GTG GAA ATC AAA GCA

IM9 Light Protein SEQ I.D. No. 11

DIQ MTQ FPSTLSASVGDRVTITCRASQSISAWLAWYQQKPGKAPKLLIY
KASSLESGVPSRFSGSGSGTEFTLTITSLQPDDFATYFCQHYNRPWTFGQGTKVEIK

SEQ. ID NO. 12

GAA ATG CAA CTG GTG GAA TTT GGG GGA GGC CTG CTA CAG CCT GGC
AGG GCC CTG AGA CTC 60
TCC TGT GCA GCC TCT GGA TTC AGG TTT GAT GAT TAT GCC ATG CAC
TGG GTC CGG CAA ACT 120
CCA GGG AAG GGC CTG GAG TGG GTC GCA GGT ATT AGT TGG AAT AGT
GAC ACC ATA GAC TAT 180
GCG GAC TCT GTG AAG GGC CGA TTC ACC ATC TCC AGA GAC AAC GCC
AAG AAC TCC CTC TAT 240
TTG CAA ATG AAC AGT CTC AGA GCT GAG GAC ACG GCC TTG TAT TAC
TGT ACA AAA AGA AGG 300
GGG GTG ACA GAC ATT GAC CCT TTT GAT ATC TGG GGC CAA GGG ACA
ATG GTC ATC GTC TCT 360
TCA GAG 366

IM9 HEAVY PROTEIN SEQ I.D. No. 13

EMQLVEFGGGLLQPGRALRLSCAASGFRFDDYAMHWVRQTPGKGLEWVAGISWNSDTIDYA
DSVKGRFTISRDNAKNSLYLQMNSLRAEDTALYYCTKRRGVTDIDPFDIWGQGTMVIVSS

SEQ I.D. No 14

DIVMTQSPSSLSVSAGERVTMSCKSSQSLLNSGNQKNFLAWYQQKPGQPPKLLIYGASTRE
SGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCQNDHSYPLTFGAGTKL

SEQ I.D. No 15

DVVMTQTPLSLPVSLGDQASISCRSSQSLVHSQGNTYLRWYLQKPGQSPKVLIYKVSNRFS
GVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTHVPWTFGGGTKLE

SEQ I.D. No 16

DIVLTQSPAIMSASPGEKVTMTCSASSSVNYMYWYQQKSGTSPKRWIYDTSKLASGVPVRF
SGSGSGTSYSLTISSMETEDAAEYYCQQWGRNPTFGGGTKLEIK

SEQ I.D. No 17

DIQMTQSPASLSASVGETVTITCRASGNIHNYLAWYQQKQGKSPQLLVYYTTTLADGVPSR
FSGSGSGTQYSLKINSLQPEDFGSYYCQHFWSTPRTFGGGTKLEIK

SEQ I.D. No 18

EIVLTQSPAITAASLGQKVTITCSASSSVSSLHWYQQKSGTSPKPWIYEISKLASGVPARF
SGSGSGTSYSLTINTMEAEDAAIYYCQQWTYPLITFGAGTKLELK

SEQ I.D. No 19

DIQMTQIPSSLSASLGDRVSISCRASQDINNFLNWYQQKPDGTIKLLIYFTSRSQSGVPSR
FSGSGSGTDYSLTISNLEQEDIATYFCQQGNALPRTFGGGTKLEIK

SEQ I.D. No 20

SVLTQPPSVSGAPGQRVTISCTGSSSNIGAGNHVKWYQQLPGTAPKLLIFHNNARFSVSKS
GSSATLAITGLQAEDEADYYCQSYDRSLRVFGGGTKLTVL

SEQ I.D. No 21

QSVLTQPPSASGTPGQRVTISCSGTSSNIGSSTVNWYQQLPGMAPKLLIYRDAMRPSGVPD
RFSGSKSGASASLAIGGLQSEDETDYYCAAWDVSLNAYVFGTGTKVTVL

SEQ I.D. No 22

EVKLVESGGGLVQPGGSLRLSCATSGFTFSDFYMEWVRQPPGKRLEWIAASRNKGNKYTTE
YSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCARNYYGSTWYFDVWGAGTTVTVSS

SEQ I.D. No 23

EVKLDETGGGLVQPGRPMKLSCVASGFTFSDYWMNWVRQSPEKGLEWVAQIRNKPYNYETY
YSDSVKGRFTISRDDSKSSVYLQMNNLRVEDMGIYYCTGSYYGMDYWGQGTSVTVSS

SEQ I.D. No 24

VQLQQSGAELMKPGASVKISCKASGYTFSDYWIEWVKQRPGHGLEWIGEILPGSGSTNYHE
RFKGKATFTADTSSSTAYMQLNSLTSEDSGVYYCLHGNYDFDGWGQGTTLTVSS

SEQ I.D. No 25

QVQLKESGPGLVAPSQSLSITCTVSGFSLTGYGVNWVRQPPGKGLEWLGMIWGDGNTDYNS
ALKSRLSISKDNSKSQVFLKMNSLHTDDTARYYCARERDYRLDYWGQGTTLTVSS

SEQ I.D. No 26

EVKLLESGGGLVQPGGSLKLSCAASGFDFSKYWMSWVRQAPGKGLEWIGEIHPDSGTINYT
PSLKDKFIISRDNAKNSLYLQMSQVRSEDTALYYCARLHYYGYNAYWGQGTLVTVSA

SEQ I.D. No. 27

EVQLQQSGVELVRAGSSVKMSCKASGYTFTSNGINWVKQRPGQGLEWIGYNNPGNGYIAYN
EKFKGKTTLTVDKSSSTAYMQLRSLTSEDSAVYFCARSEYYGGSYKFDYWGQGTTLTVSS

SEQ I.D. No 28

VKLEQSGPGLVRPSQTLSLTCTVSGTSFDDYYSTWVRQPPGRGLEWIGYVFYHGTSDTDTP
LRSRVTMLVNTSKNQFSLRLSSVTAADTAVYYCARNLIAGCIDVWGQGSLVTVSS

SEQ I.D. No 29

EVQLVQSGGGVVQPGRSLRLSCSSSGFIFSSYAMYWVRQAPGKGLEWVAIIWDDGSDQHYA
DSVKGRFTISRNDSKNTLFLQMDSLRPEDTGVYFCARDGGHGFCSSASCFGPDYWGQGTPV
TVSS

DIQMTQFPST LSASVGDRVN ITCRASGFTF SNFGMHWIRQ KPGKGLKWVA
YISGGSSTVH YADSLKGRFT ISRDNPKNEL FLTITSLQPD DFAMYYCARD
YYVNNYWYFD VWGQGTKVEI KR (122 residues) **(SEQ. I.D. NO. 30)**

SEQ. I.D. Nos. 31 - TAGTGGATCCAACTGATTTCTCCAT
SEQ. I.D. No. 32 - TTATTTACTTCTGGGTCACCAGGTTTATTC
SEQ. I.D. No. 33 - AAGAGGCCGAGCTGGCCCTTCCCTGAATAACCAGGCAGT
SEQ. I.D. No. 34 - GGGAAGGGCCAGCTCGGCGTGTTCCTATAATATGATCAA
SEQ. I.D. No. 35 - TTCCTGGCCCTGCAGGCCCAGTTGTCTGTGTCTTCTGTT
SEQ. I.D. No. 36 - AACTGGGCCTGCAGGGCCAGGAAGCAAAGTTTAAATTCTA
SEQ. I.D. No. 37 - CATGTCTGGATCCAACTGATTT
SEQ. I.D. No. 38 - CTGATTTACTTCTGGGTGACCAGGTTTATTCAA

SEQ. I.D. No. 39

5'-AAGGGCCAGCTCGGCCTCTTCCTATAATATGATCAATAGTATAAATATTTGTGTTTC-
TATTTCCAATCTCAGGTGCCAAATGTGACATCCAGATGACCCA-3'

SEQ. I.D. No. 40

5'-
TGGGCCTGCAGGGCCAGGAAGCAAAGTTTAAATTCTACTCACGTTTGATTTCCACCTTGG-
TT-3'

#1 = B695 = 5' -GGG-AAG-GGC-CAG-CTC-GGC-CTC-TTC-CTA-TAA-TAT-
GAT-CAA-TAG-TAT-AAA-TAT-TTG-TGT-TTC-TAT-TTC-CAA-TCT-CAG-GTG-
CCA-AAT-GTG-ACA-TCC-AGA-TGA-CCC-AGT-TTC-CT- 3 **(SEQ. I.D. NO. 41)**

#2 = B696 = 5' -GCA-TGC-CGA-AGT-TGG-AGA-AGG-TGA-AGC-CGG-AGG-
CGC-GGC-AGG-TGA-TGT-TCA-CGC-GGT-CGC-CCA-CGG-AGG-CGG-ACA-GGG-
TGG-AAG-GAA-ACT-GGG-TCA-TCT-GGA-TGT- 3 **(SEQ. I.D. NO. 42)**

B549 = 5' -GGC-TTC-ACC-TTC-TCC-AAC-TTC-GGC-ATG-CAC-TGG-ATC-CGC-CAG-AAG-CCC-GGC-AAG-GGC-CTG-AAG-TGG-GTG-GCC-TAC-ATC-TCC-GGC-GGC-TCC-TCC-ACC-GTG-CAC-TA- 3 **(SEQ. I.D. NO. 43)'**

B550 = 5' -GGT-GAT-GGT-CAG-GAA-CAG-CTC-GTT-CTT-GGG-GTT-GTC-GCG-GGA-GAT-GGT-GAA-GCG-GCC-CTT-CAG-GGA-GTC-GGC-GTA-GTG-CAC-GGT-GGA-GGA-GCC-GCC-GGA-GAT-GTA- -3 **(SEQ. I.D. NO. 44)'**

B697 = 5' -CCC-CAA-GAA-CGA-GCT-GTT-CCT-GAC-CAT-CAC-CTC-CCT-GCA-GCC-CGA-CGA-CTT-CGC-CAT-GTA-CTA-CTG-CGC-CCG-CGA-CTA-CTA-CGT-GAA-CAA-CTA-CTG-GTA-CTT-CGA-CGT-GT **(SEQ. I.D. NO. 45)**

SEQ. I. D. No. 46 #6 = B698 = 5' -CAC-AGA-CAA-CTG-GGC-CTG-CAG-GGC-CAG-GAA-GCA-AAG-TTT-AAA-TTC-TAC-TCA-CGT-TTTG-ATC-TCC-ACC-TTG-GTG-CCC-TGG-CCC-CAC-ACG-TCG-AAG-TAC-CAG-TAG-TT

SEQ. I.D. No. 47 - 5' -GGG-AAG-GGC-CAG-CTC-GGC-CTC-TT -3'

SEQ. I.D. No. 48 - 5'-CAC-AGA-CAA-CTG-GGC-CTG-CA- 3'

SEQ I.D. No. 49 - 5'-CAC-CAT CCT GTT TGC TTC TTT CCT CAG GAA CTG TGC ACT GGC ACC ACC ACC CAT AGA GGG AGA AGT GCC CCC ACC TGC TCC TCA GTT -3'

SEQ. I.D. No. 50

5'-GGGTAAAAATAGAATGAAGGATGATTTTTATAAAT-3'

SEQ. I.D. No. 51 GAC ATC CAG ATG ACC CAG TTT CCT TCC ACC CTG TCC GCC TCC GTG GGC GAC CGC GTG AAC ATC ACC TGC CGC GCC TCC GGC TCC ACC TTC TCC AAC TTC GGC ATG CAC TGG ATC CGC CAG AAG CCC GGC AAG GGC CTG AAG TGG GTG GCC TAC ATC TCC GGC GGC TCC TCC ACC GTG CAC TAC GCC AAC TCC CTG AAG GGC CGC TTC ACC ATC TCC CGC GAC AAC CCC AAG AAC GAG CTG TTC CTG ACC ATC ACC TCC CTG CAG CCC GAC GAC TTC GCC ATG TAC TAC TGC GCC CGC GAC TAC TAC GTG AAC AAC TAC TGG TAC TTC GAC GTG TGG GGC CAA GGG ACC AAG GTG GAA ATC AAA

5'-AAG-AGC-TCC-TGA-ACC-TCG-CGG-ACA-GTT-AA-3' **SEQ. I.D. No. 52**

5'-AAA-TCG-ATC-TCA-GGC-CTC-AGA-CTC-GGC-CTG-ACC-CGT-GGA-AA-3' **SEQ. I.D. No. 53**

SEQ. I.D. No. 54 5'-CTG TTT GCT TCT TTC CTC AGG AGG CGG TTC AGG AGG ATC AGG CGG TTC AGG TGG ATC AGG AGG CGA CAT CCA GAT GAC CCA GTC TCC T-3'

SEQ. I.D. No. 55 - 5'-GTC AGG CTG GAA CTG AGG AGC AGG TGG GGG CAC TTC TCC CTC TAT GGG TGA TGG TGC CAA TGT TTG ATT TCC ACC TTG GTC CCT TGG CCG -AA-3'

SEQ. I.D. No. 56 - 5'-GAG AAG TGC CCC CAC CTG CTC CTC AGT TCC AGC CTG ACC CCC TCC CAT CCT -3'

SEQ I.D. No. 59

GAA ATG CAA CTG GTG GAA TCT GGG GGA GGC CTG CTA CAG CCT GGC CGG GCC CTG CGG CTC TCC TGT GCA GCC TCT GGA TTC ACT TTT AGT AAC TTT GGA ATG CAC TGG ATT CGG CAA ACT CCA GGG AAG GGC CTG GAG TGG GTC GCA TAC ATT AGT GGT GGC AGT AGT ACC GTC CAC TAT GCA GAC TCC TTG AAG GGC CGA TTC ACC ATC TCC CGG GAC AAC GCC AAG AAC TCC CTC TAT TTG CAA ATG ACC AGT CTC CGG GCT GAG GAC ACG GCC TTG TAT TAC TGT GCA CGG GAT TAC TAC GTT AAT AAC TAC TGG TAC TTC GAT GTC TGG GGC CAA GGG ACA ATG GTC ATC GTC TCT TCA G

SEQ I.D. No. 60

5'-GAT CCG AAA TGC AAC TGG TGG AAT CTG GGG GAG GCC TGC TAC AGC CTG GCC GGG CCC TGC GGC TCT CCT GTG CAG CCT CTG GAT TCA CCT TTA G-3'

SEQ I.D. No. 61

5'-CAC CAC TAA TGT ATG CGA CCC ACT CCA GGC CCT TCC CTG GAG TTT GCC GAA TCC AGT GCA TTC CAA AGT TAC TAA AGG TGA ATC CAG AGG C-3'

SEQ I.D. No. 62

5'-GGG TCG CAT ACA TTA GTG GTG GCA GTA GTA CCG TCC ACT ATG CAG ACT CCT TGA AGG GCC GAT TCA CCA TCT CCC GGG ACA ACG CCA AGA A 3'

SEQ I.D. No. 63

5'-TAT TAC TGT GCA CGG GAT TAC TAC GTT AAT AAC TAC TGG TAC TTC GAT GTC TGG GGC CCA GGG ACA ATG GTC ATC GTC TCT TCA -3'

SEQ I.D. No. 64

5'-GTA ATC CCG TGC ACA GTA ATA CAA GGC CGT GTC CTC AGC CCG GAG ACT GTT CAT TTG CAA ATA GAG GGA GTT CTT GGC GTT GTC CCG GGA G -3'

SEQ I.D. No. 65

5'-AAG GAT CCG AAA TGC AAC TGG TGG AAT CT -3'

SEQ I.D. No. 66 - GAC GAA TTC TGA AGA GAC GAT GAC CAT TG

We Claim:

1. A recombinant antibody or antigen binding fragment thereof, comprised of at least one light chain variable domain, which domain, in turn, comprises three CDRs wherein the amino acid sequence of one or more of the CDRs is derived from the amino acid sequence of the corresponding CDR(s) of a heavy chain variable domain of one (donor) antibody and further comprises four framework regions wherein the amino acid sequence of one or more of the framework regions is derived from the amino acid sequence of the corresponding framework region(s) from the light chain variable domain of the same or a different (acceptor) antibody.

2. A recombinant antibody or antigen binding fragment thereof of claim 1, wherein the antibody or antigen binding fragment thereof is selected from the group consisting of:

a) a CSV_L fragment;

b) a heavy body [CSV_L--C_L];

c) a kappa body fragment [CDR-grafted V_L--C_L || CSV_L--C_L];

d) an intact kappa body (2X [CDR-grafted V_L--C_L || CSV_L--C_H]); or

e) an ScFv(CSV_L) fragment [either CDR-grafted V_L-linker--CSV_L or CSV_L --linker-- CDR-grafted V_L.]

3. A recombinant fragment of claim 2, wherein the donor and acceptor antibodies are independently chosen from the group consisting of murine, rabbit, and primate antibodies.

4. A recombinant antibody fragment of claim 3, wherein the amino acid sequences of all three CDRs of the

CSVL domain are derived from the amino acid sequences of the corresponding CDRs of the heavy chain variable domain of the donor antibody and the amino acid sequences of all four framework regions of the CSVL domain are derived from the amino acid sequences of the corresponding framework regions of the light chain variable domain of the acceptor antibody.

5. A recombinant antibody fragment of claim 4, wherein the acceptor antibody is human.

6. A recombinant antibody fragment of claim 5, wherein the human acceptor antibody has light chains of the kappa class.

7. A recombinant antibody fragment of claim 6, wherein the acceptor antibody is of the IgG class.

8. A recombinant antibody fragment of claim 7, wherein the donor antibody is murine.

9. A recombinant antibody fragment of claim 8, wherein the donor murine antibody has affinity for tumor antigens or antigens on thrombi.

10. A recombinant antibody fragment of claim 9, wherein the donor murine antibody fragment has affinity for tumor antigens.

11. A recombinant antibody fragment of claim 10, wherein the donor murine antibody has affinity for the tumor markers chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear Antigen, Vimentin,

and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes.

12. A recombinant antibody fragment of claim 11, wherein the donor murine antibody has affinity for the tumor antigen CEA.

13. A recombinant antibody fragment of claim 12, wherein the donor antibody is ZCE 025 or CEM 231.

14. A recombinant antibody fragment of claim 13, wherein the acceptor antibody is IM9, and the framework regions are mostly the same as the corresponding IM9 framework regions.

15. A recombinant antibody fragment of claim 14, wherein the donor antibody is ZCE 025.

16. A recombinant antibody fragment of claim 2, wherein the C-terminus or N-terminus of the fragment molecule is fused to a metal chelating peptide.

17. A recombinant antibody fragment of claim 16, wherein the metal chelating peptide has the amino acid sequence HWHHHP (**Sequence I.D.No. 2**) and is fused to the C-terminus of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

18. A recombinant antibody fragment of claim 4, wherein the fragment is selected from the group consisting of a) kappabody fragment and b) intact kappabody, the amino acid sequences of all three CDRs of the CDR-Grafted V_L domain are derived from the amino acid sequences of the corresponding CDRs of the V_L domain of the CDR-Grafted donor antibody, the amino acid sequences of all four framework regions of the CDR-Grafted V_L domain are derived from the amino acid sequences of the corresponding framework regions of the V_L

domain of the CDR-Grafted acceptor antibody, and the light chain constant domains are identical in sequence to the corresponding constant domains of the acceptor antibody or antibodies.

19. A recombinant antibody fragment of claim 18, wherein the amino acid sequence of the framework regions of the CSV_L and the CDR-Grafted V_L and the C_L domains are derived from the same human acceptor antibody.

20. A recombinant antibody fragment of claim 19, wherein the framework regions of the CSV_L and the CDR-grafted V_L, as well as the complete C_L domains are derived from the corresponding regions and domains of a human acceptor antibody whose light chain is of the kappa class.

21. A recombinant antibody fragment of claim 20, wherein the acceptor antibody is of the IgG class.

22. A recombinant antibody fragment of claim 21, wherein the donor antibody for both the CSV_L and the CDR-grafted V_L is the same murine antibody.

23. A recombinant antibody fragment of claim 22, wherein the donor murine antibody has affinity for a tumor antigen or an antigen on thrombi.

24. A recombinant antibody fragment of claim 23, wherein the donor murine antibody has affinity for a tumor antigen.

25. A recombinant antibody fragment of claim 24, wherein both donor murine antibodies have affinity for tumor antigens chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen

Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes.

26. A recombinant antibody fragment of claim 25, wherein the donor murine antibody has affinity for the tumor antigen CEA.

27. A recombinant antibody fragment of claim 26, wherein the donor murine antibody is either ZCE 025 or CEM 231.1.

28. A recombinant antibody fragment of claim 27, wherein the human acceptor antibody is IM9, and the amino acid sequences of both sets of framework regions are derived from the amino acid sequences of the corresponding IM9 light chain framework regions.

29. A recombinant antibody fragment of claim 28, wherein the murine donor antibody is ZCE 025.

30. A recombinant antibody fragment of claim 29, wherein the C-terminus or the N-terminus of either the CSV_L -containing or the CDR-Grafted - containing chain of the fragment molecule is fused to a metal chelating peptide.

31. A recombinant antibody fragment of claim 30, wherein the metal chelating peptide has the amino acid sequence HWHHHP and is fused to the C-terminus of the CSV_L containing chain of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

32. A recombinant antibody or fragment thereof of claim 2, wherein the antibody or fragment thereof is the fragment ScFv(CSV_L), wherein a CDR-Grafted V_L domain is

covalently bonded to a CSV_L domain through a polypeptide linker.

33. A recombinant fragment of claim 32, wherein the donor and acceptor antibodies are independently chosen from the group consisting of murine, rabbit, and primate antibodies.

34. A recombinant antibody fragment of claim 33, wherein

a) The amino acid sequence of all three CDRs of the CSV_L domain derived from those of the corresponding CDRs of the heavy chain variable domain of the donor antibody used;

b) The amino acid sequence of all four CSV_L framework regions are derived from those of the corresponding framework regions of the light chain variable domain of the acceptor antibody used;

c) The amino acid sequences of all three CDRs of the CDR-Grafted V_L domain are derived from those of the corresponding CDRs of the V_L domain of the donor antibody used; and

d) The amino acid sequences of all four framework regions of the CDR-Grafted V_L domain are derived from those of the corresponding framework regions of the V_L domain of the acceptor antibody.

35. A recombinant antibody fragment of claim 34, wherein the amino acid sequences of the framework regions of both the CSV_L and the CDR-Grafted V_L are derived from those of the corresponding framework regions of the same human acceptor antibody.

36. A recombinant antibody fragment of claim 35, wherein the polypeptide linker is composed of about 12 to about 18 amino acids.

37. A recombinant antibody fragment of claim 36, wherein the C-terminus of the CDR-Grafted V_L domain is fused to the N-terminus of the polypeptide linker, and wherein the C-terminus of the polypeptide linker is bonded to the N-terminus of the CSV_L domain.

38. A recombinant antibody fragment of claim 37, wherein the donor antibody is murine.

39. A recombinant antibody fragment of claim 38, wherein the donor murine antibody has affinity for tumor antigens or antigens on thrombi.

40. A recombinant antibody fragment of claim 39, wherein the donor murine antibody has affinity for a tumor antigen.

41. A recombinant antibody fragment of claim 40, wherein the donor murine antibody has affinity for tumor antigens chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes

42. A recombinant antibody fragment of claim 41, wherein the donor murine antibody has affinity for the tumor antigen CEA.

43. A recombinant antibody fragment of claim 42, wherein the linker polypeptide is composed of serine and glycine amino acid residues.

44. A recombinant antibody fragment of claim 43, wherein the donor murine antibody is either ZCE 025 or CEM 231.1.

45. A recombinant antibody fragment of claim 44, wherein the human acceptor antibody is IM9, and both sets of framework regions are mostly the same as the corresponding IM9 light chain framework regions.

46. A recombinant antibody fragment of claim 45, wherein the murine donor antibody is ZCE025.

47. A recombinant antibody fragment of claim 46, wherein the linker polypeptide is of the formula -GGSGGSGGSGGSGG-.

48. A recombinant antibody fragment of claim 47, wherein the C-terminus or the N-terminus of the ScFv(C_SV_L) is fused to a metal chelating peptide.

49. A recombinant antibody fragment of claim 48 wherein the metal chelating peptide has the amino acid sequence HWHHHP and is fused to the C-terminus of the CSV_L domain of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

50. A DNA or RNA sequence coding for a recombinant antibody or fragment thereof, wherein the antibody or fragment thereof is comprised of at least one light chain variable domain, which domain, in turn, comprises three CDRs wherein the amino acid sequence of one or more of the CDRs is derived from the amino acid sequence of the corresponding

CDR(s) of a heavy chain variable domain of one (donor) antibody and further comprises four framework regions wherein the amino acid sequence of one or more of the framework regions are derived from the amino acid sequence of the corresponding framework region(s) from the light chain variable domain of the same or a different (acceptor) antibody.

51. A DNA or RNA sequence coding for a recombinant antibody or antigen binding fragment thereof of claim **50**, wherein the recombinant antibody or antigen binding fragment thereof is selected from the group consisting of:

a) a CSV_L fragment;
b) a heavy body [CSV_L--C_L];
c) a kappa body fragment [CDR-grafted V_L--C_L || CSV_L--C_L];
d) an intact kappa body (2X [CDR-grafted V_L--C_L || CSV_L--C_H]); or
e) an ScFv(CSV_L) fragment [either CDR-grafted V_L--linker--CSV_L or CSV_L --linker-- CDR-grafted V_L.]

52. A DNA or RNA sequence of claim **51**, wherein in the recombinant fragment the donor and acceptor antibodies that are coded for are independently chosen from the group consisting of murine, rabbit, and primate antibodies.

53. A DNA or RNA sequence of claim **52**, wherein the amino acid sequences of all three CDRs of the CSVL domain that are coded for are derived from the amino acid sequences of the corresponding CDRs of the heavy chain variable domain of the donor antibody and the amino acid sequences of all four framework regions of the CSVL are derived from the amino acid sequences of the corresponding framework regions of the light chain variable domain of the acceptor antibody.

54. A DNA or RNA sequence of claim 53, wherein the acceptor antibody that is coded for is human.

55. A DNA or RNA sequence of claim 54, wherein the human acceptor antibody that is coded for has light chains of the kappa class.

56. A DNA or RNA sequence of claim 55, wherein the acceptor antibody that is coded for is of the IgG class.

57. A DNA or RNA sequence of claim 56, wherein the donor antibody that is coded for is murine.

58. A DNA or RNA sequence of claim 57, wherein the donor murine antibody that is coded for has affinity for tumor antigens or antigens on thrombi.

59. A DNA or RNA sequence of claim 58, wherein the donor murine antibody fragment that is coded for has affinity for tumor antigens.

60. A DNA or RNA sequence of claim 59, wherein the donor murine antibody that is coded for has affinity for the tumor markers chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes.

61. A DNA or RNA sequence of claim 60, wherein the donor murine antibody that is coded for has affinity for the tumor antigen CEA.

62. A DNA or RNA sequence of claim 61, wherein the donor antibody that is coded for is ZCE 025 or CEM 231.

63. A DNA or RNA sequence of claim 62, wherein the acceptor antibody that is coded for is IM9, and the framework regions are mostly the same as the corresponding IM9 framework regions.

64. A DNA or RNA sequence of claim 63, wherein the donor antibody that is coded for is ZCE 025.

65. A DNA or RNA sequence of claim 51 wherein the 5'-terminus or 3'-terminus of the DNA or RNA coding for the fragment molecule is fused to a DNA or RNA sequence respectively, coding for metal chelating peptide.

66. A DNA or RNA sequence of claim 65, wherein the metal chelating peptide that is coded for has the amino acid sequence HWHHHP (Sequence I.D.No.) and is fused to the C-terminus of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

67. A DNA or RNA sequence of claim 51, whereinthe fragment encoded is selected from the group consisting of a) kappabody and b) intact kappabody and the amino acid sequences of all three CDRs that are coded for of the CDR-Grafted V_L domain are derived from the amino acid sequences of the corresponding CDRs of the V_L domain of the CDR-Grafted donor antibody and the amino acid sequences of all four framework regions that are coded for of the CDR-Grafted V_L domain are derived from the amino acid sequences of the corresponding framework regions of the V_L domain of the CDR-Grafted acceptor antibody, and the light chain constant domains that are coded for are identical in sequence to the corresponding constant domains of the acceptor antibody or antibodies.

68. A DNA or RNA sequence of claim 67, wherein the amino acid sequences of the framework regions of both the CSV_L and the CDR-Grafted V_L and the amino acid sequence of the C_L domains coded for acceptor antibodies are derived from the same human acceptor antibody.

69. A DNA or RNA sequence of claim 68, wherein the amino acid sequence of the framework regions of the CSV_L and the CDR-Grafted V_L and the amino acid sequence of the C_L domains that are coded for are derived from the corresponding regions and domains of a human acceptor antibody whose light chain is of the kappa class.

70. A DNA or RNA sequence of claim 69, wherein the acceptor antibody that is used is of the IgG class.

71. A DNA or RNA sequence of claim 70, wherein the donor antibody used for both is the same murine antibody.

72. A DNA or RNA sequence of claim 71, wherein the donor murine antibody used has affinity for a tumor antigen or an antigen on thrombi.

73. A DNA or RNA sequence of claim 72, wherein the donor murine antibody used has affinity for a tumor antigen.

74. A DNA or RNA sequence of claim 73, wherein the donor murine antibody that is used has affinity for tumor antigens chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear

Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes.

75. A DNA or RNA sequence of claim 74, wherein the donor murine antibody used has affinity for the tumor antigen CEA.

76. A DNA or RNA sequence of claim 75, wherein the donor murine antibody used is either ZCE 025 or CEM 231.1.

77. A DNA or RNA sequence of claim 76, wherein the human acceptor antibody used is IM9, and both sets of framework regions are derived from the corresponding IM9 light chain framework regions.

78. A DNA or RNA sequence of claim 77, wherein the murine donor antibody used is ZCE 025.

79. A DNA or RNA sequence of claim 78, wherein the C-terminus or the N-terminus of either the CSV_L -containing or the CDR-Grafted - containing chain of the fragment molecule is fused to a metal chelating peptide.

80. A DNA or RNA sequence of claim 79, wherein the metal chelating peptide that is coded for has the amino acid sequence HWHHHP and is fused to the C-terminus of the CSV_L containing chain of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

81. A DNA or RNA sequence coding for a recombinant antibody or fragment thereof of claim 51, wherein the antibody or fragment thereof that is coded for is the fragment $ScFv(CSV_L)$, wherein a CDR-Grafted V_L domain is covalently bonded to a CSV_L domain through a polypeptide linker.

82. A DNA or RNA sequence of claim 81, wherein the donor and acceptor antibodies that are coded for are

independently chosen from the group consisting of murine, rabbit, and primate antibodies.

83. A DNA or RNA sequence of claim 82, wherein

a) The amino acid sequences of all three CDRs of the CSV_L domain that are coded for are derived from those of the corresponding CDRs of the heavy chain variable domain of the donor antibody used;

b) The amino acid sequences of all four CSV_L framework regions that are coded for are derived from those of the corresponding framework regions of the light chain variable domain of the acceptor antibody used;

c) The amino acid sequences of all three CDRs of the CDR-Grafted V_L domain that are coded for are derived from those of the corresponding CDRs of the V_L domain of the donor antibody used; and

d) The amino acid sequences of all four framework regions of the CDR-Grafted V_L domain that are coded for are derived from those of the corresponding framework regions of the V_L domain of the acceptor antibody used.

84. A DNA or RNA sequence of claim 83, wherein the amino acid sequences of the framework regions of both the CSV_L and the CDR-Grafted V_L are derived from those of the corresponding framework regions of the same human acceptor antibody.

85. A DNA or RNA sequence of claim 84, wherein the polypeptide linker that is coded for is composed of about 12 to about 18 amino acids.

86. A DNA or RNA sequence of claim 85, wherein the C-terminus of the CDR-Grafted V_L domain that is coded for is fused to the N-terminus of the polypeptide linker, and wherein the C-terminus of the polypeptide linker that is coded for is fused to the N-terminus of the CSV_L domain.

87. A DNA or RNA sequence of claim 86, wherein the donor antibody that is used is murine.

88. A DNA or RNA sequence of claim 87, wherein the donor murine antibody that is used has affinity for tumor antigens or antigens on thrombi.

89. A DNA or RNA sequence of claim 88, wherein the donor murine antibody that is used has affinity for a tumor antigen.

90. A DNA or RNA sequence of claim 89, wherein the donor murine antibody that is used has affinity for tumor antigens chosen from the group consisting of AFP, CA-125, CEA, Neuron Specific Enolase, C-erb2/Her-2/NEU protein, Cathepsin D, Chromagranins A, B, and C, the Cytokeratins, Epidermal Growth Factor Receptor, Epithelial Membrane Antigen, Estrogen Receptor, Progesterone Receptor, Prostatic Acid Phosphatase, Prostate Specific Antigen, Ki-67, PGP-170 (MDR), PGP-180 (MDR), p120, Proliferating Cell Nuclear Antigen, Vimentin, and the proteins expressed by the c-myc, N-myc, N-ras, Ki-ras and Ha-ras oncogenes

91. A DNA or RNA sequence of claim 90, wherein the donor murine antibody used has affinity for the tumor antigen CEA.

92. A DNA or RNA sequence of claim 91, wherein the linker polypeptide that is coded for is composed of serine and glycine amino acid residues.

93. A DNA or RNA sequence of claim 92, wherein the donor murine antibody used is either ZCE 025 or CEM 231.1.

94. A DNA or RNA sequence of claim 93, wherein the human acceptor antibody used is IM9, and the amino acid sequences of both the CSV_L and CDR-Grafted V_L framework regions that are coded for are derived from those of the corresponding IM9 light chain framework regions.

95. A DNA or RNA sequence of claim 94, wherein the murine donor antibody used is ZCE025.

96. A DNA or RNA sequence of claim 95, wherein the linker polypeptide that is coded for is of the formula -GGSGGSGGSGGSGG-.

97. A DNA or RNA sequence of claim 96, wherein the C-terminus or the N-terminus of the $ScFr(CSV_L)$ that is coded for is fused to a metal chelating peptide.

98. A DNA or RNA sequence of claim 97 wherein the metal chelating peptide that is coded for has the amino acid sequence HWHHHP and is fused to the C-terminus of the CSV_L domain of the fragment molecule through the N-terminal histidine residue of the chelating peptide.

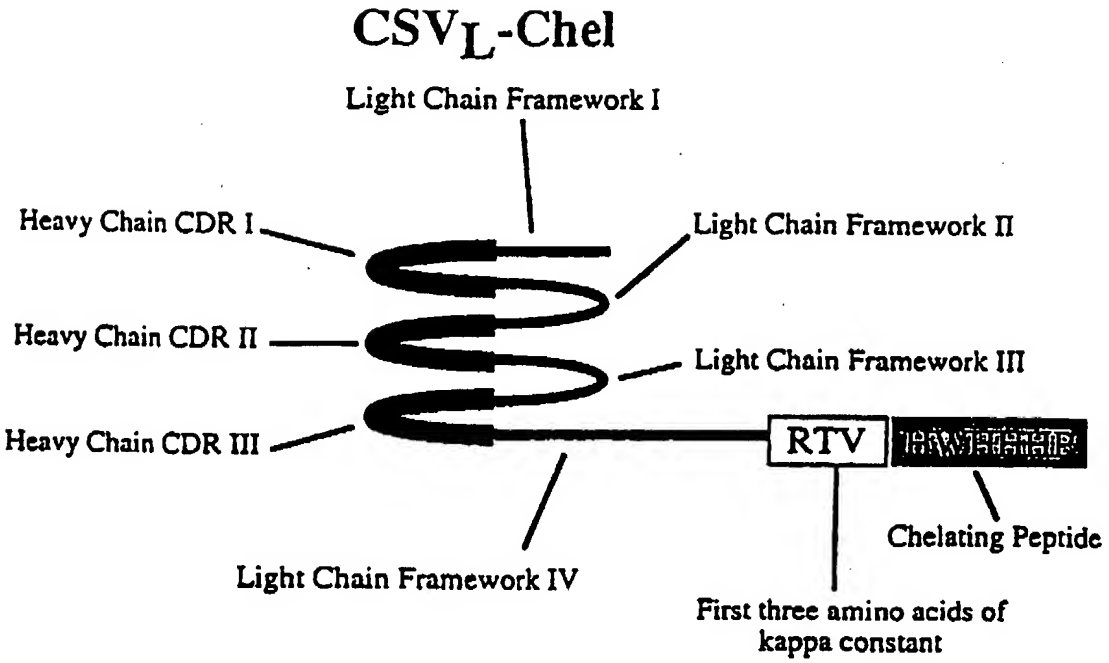

CSV$_L$-Chel
Light Chain Framework I
Heavy Chain CDR I
Light Chain Framework II
Heavy Chain CDR II
Light Chain Framework III
Heavy Chain CDR III
RTV
Chelating Peptide
Light Chain Framework IV
First three amino acids of kappa constant

FIG. 1

Heavybody

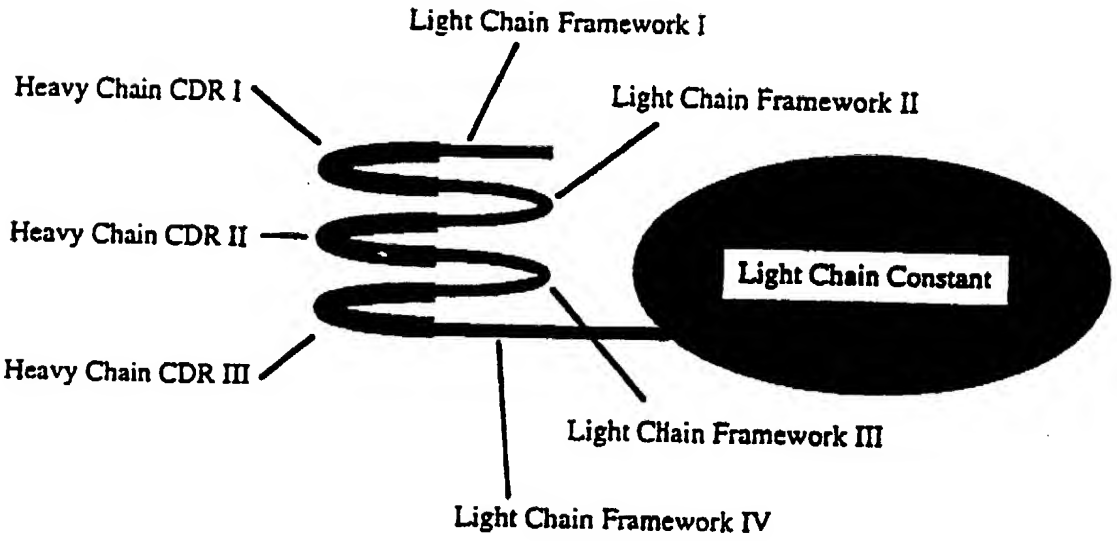


FIG. 2

Kappabody

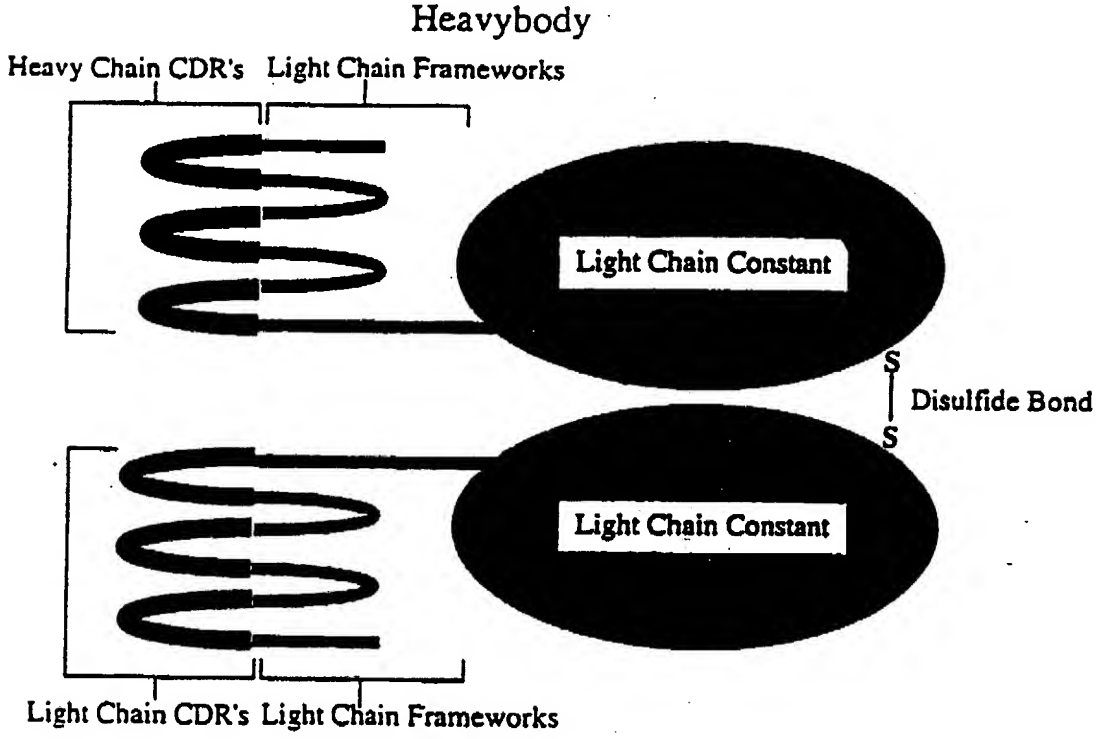


FIG. 3

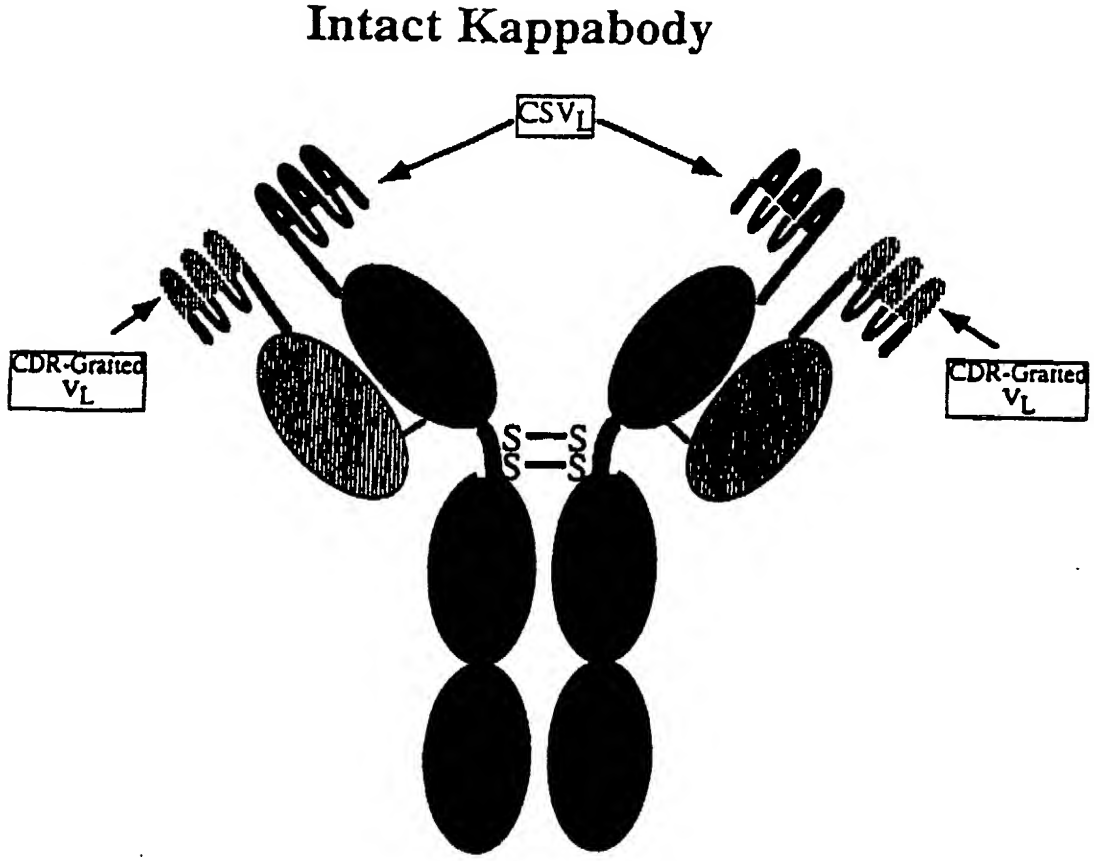


FIG. 4

SCFv-CSV$_L$

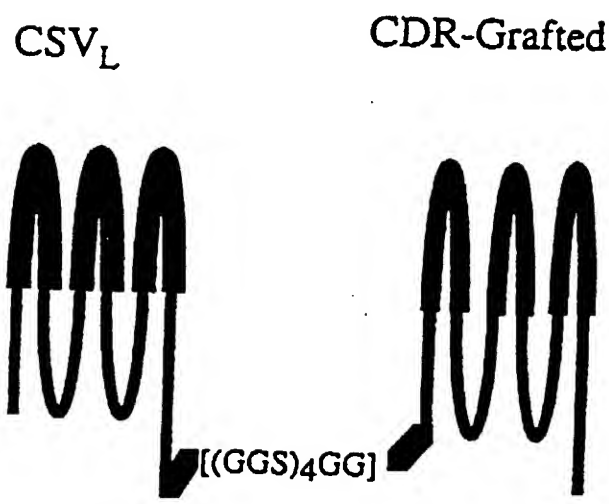


FIG. 5

FIG. 6 Alignment of immunoglobulin light chain variable domain sequences showing SCR consensus boxes.

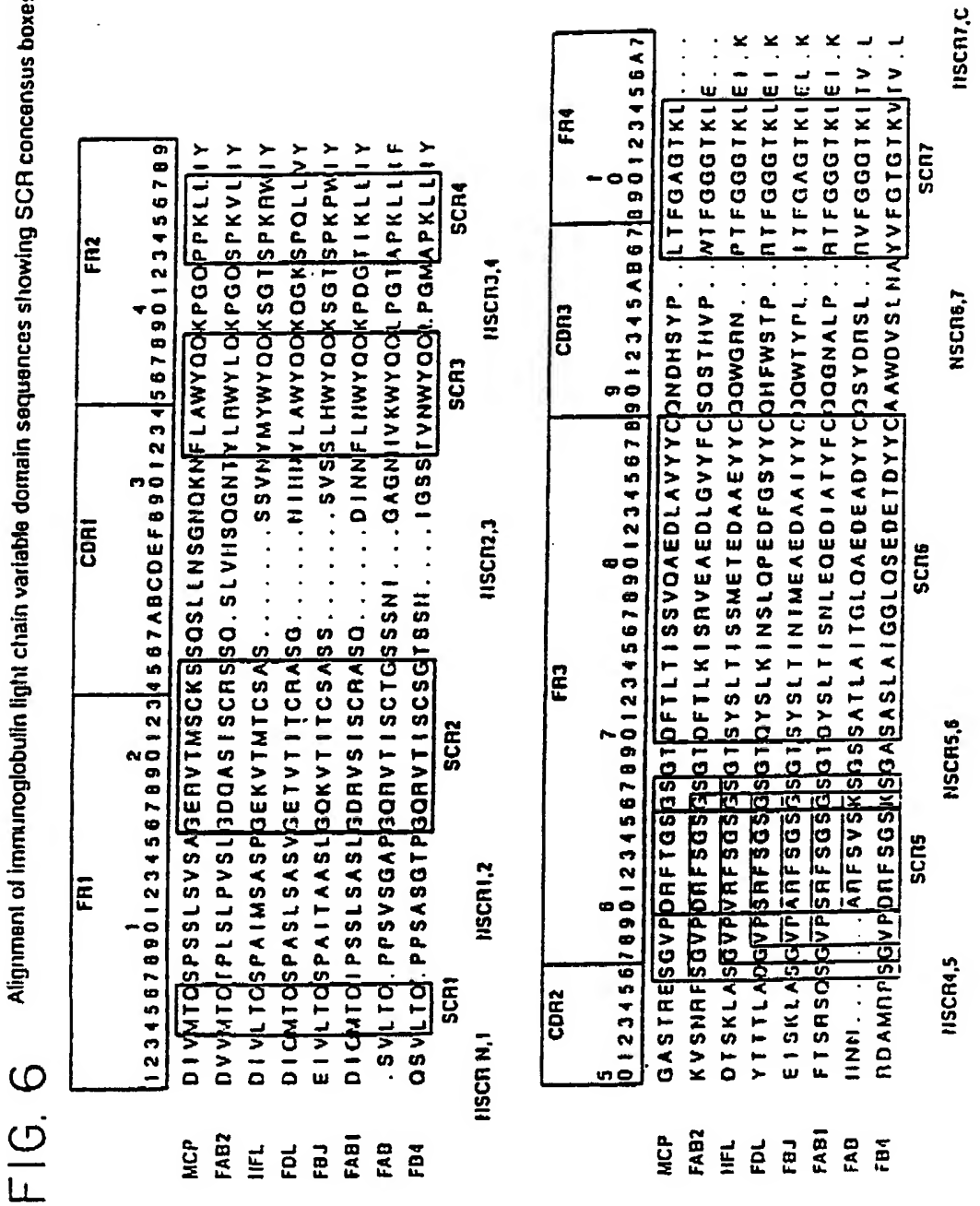


FIG. 7 Alignment of immunoglobulin heavy chain variable domain sequences showing SCR concensus boxes.

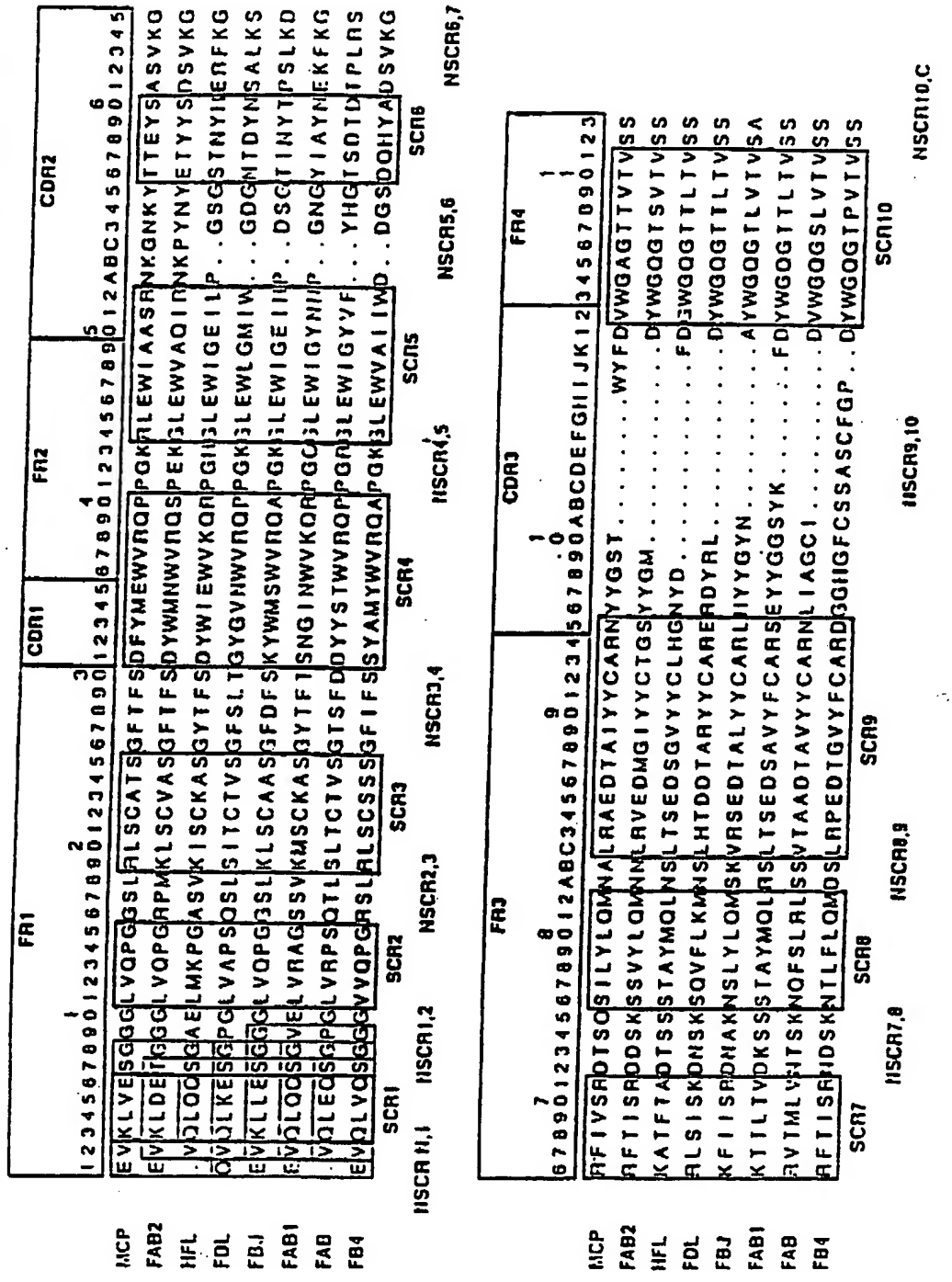
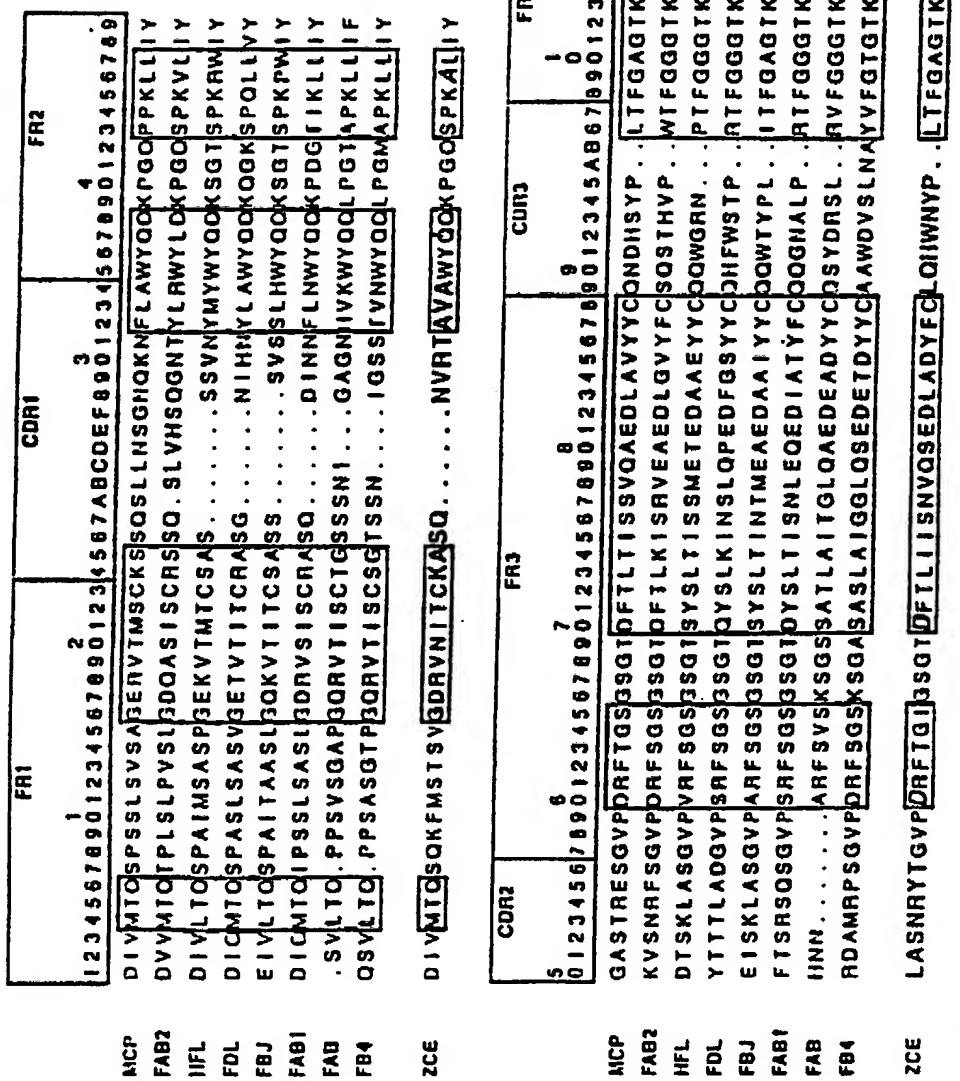


FIG. 8

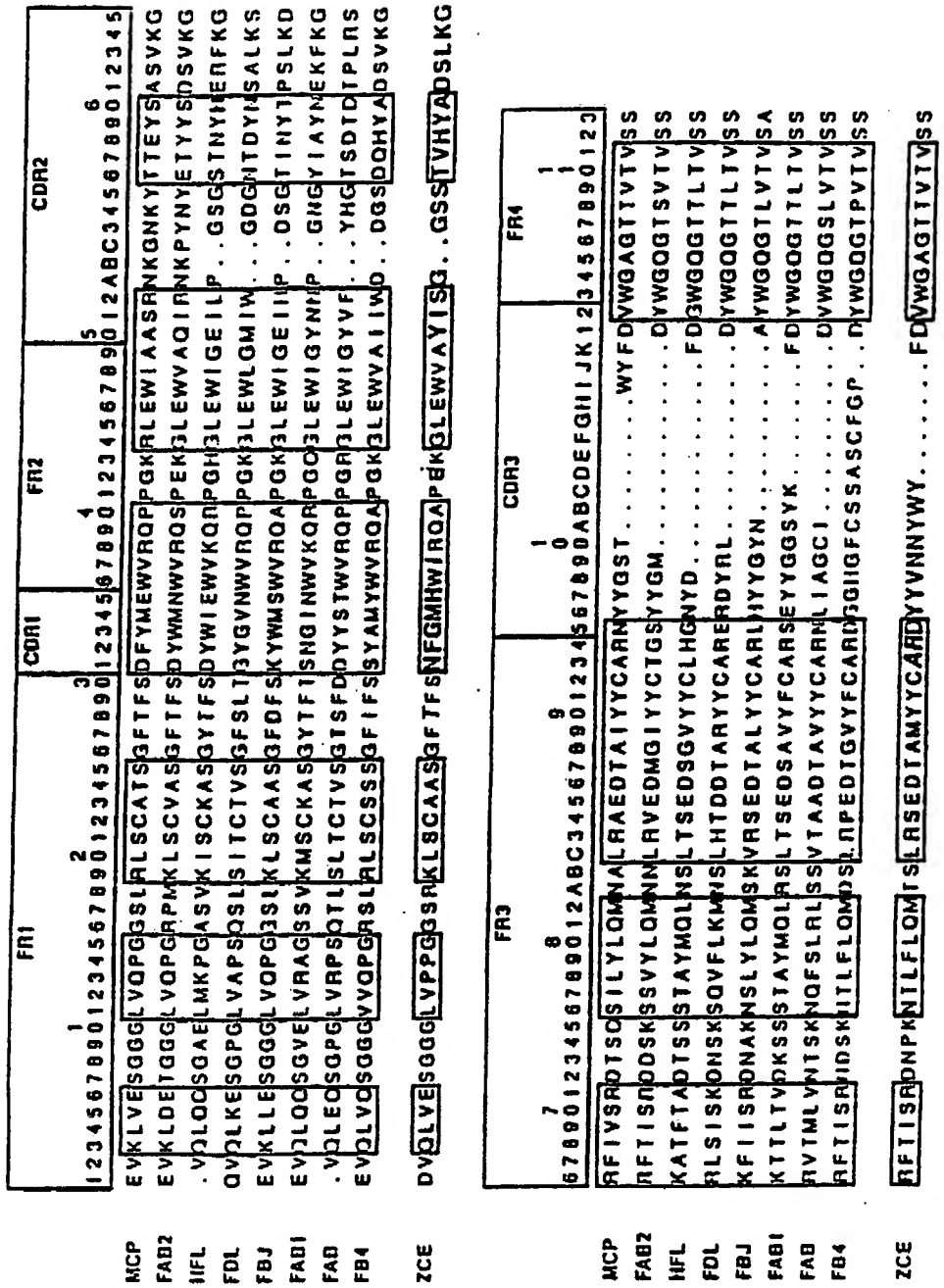
Alignment of ZCE025 light chain variable domain sequence with light chain variable domain sequences from immunoglobulins with known three-dimensional structure, showing SCR's, CDR's and CDR associated residues.



Bold residues are Kabat-defined CDR residues.
Bold italic residues are CDR associated residues.

FIG. 9

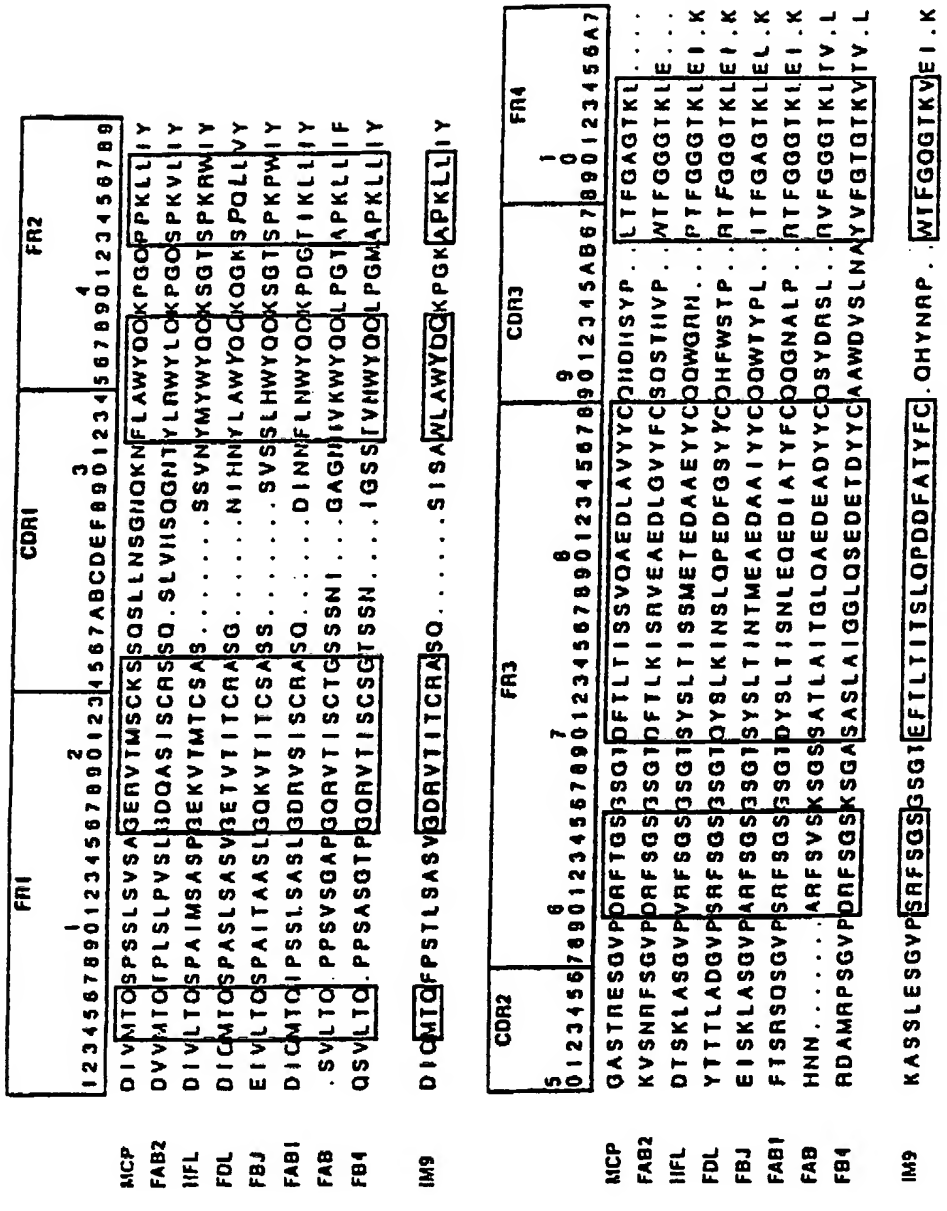
Alignment of ZCE025 heavy chain variable domain sequence with heavy chain variable domain sequences from immunoglobulins with known three-dimensional structure, showing SCR's, CDR's and CDR associated residues.



Bold residues are Kabat-defined CDR residues.
Bold italic residues are CDR associated residues.

FIG. 10

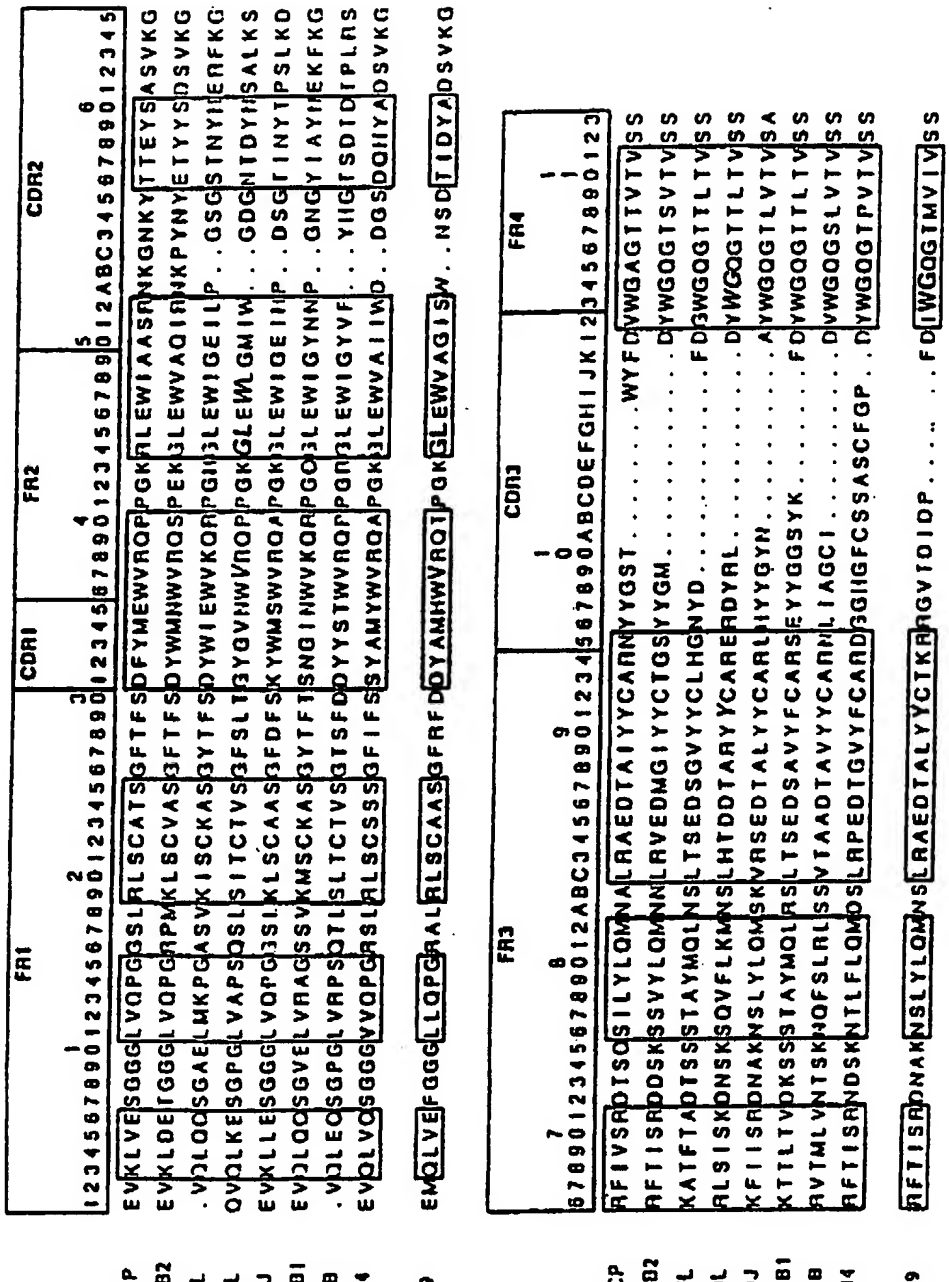
Alignment of IM9 light chain variable domain sequence with light chain variable domain sequences from immunoglobulins with known three-dimensional structure, showing SCR's and chain association residues.



Chain association residues identified in the known structure selected as the template for IM9 chain association are italicized bold. Residues identified in the IM9 Fv model as chain association residues are bold in the IM9 sequence.

FIG. 11

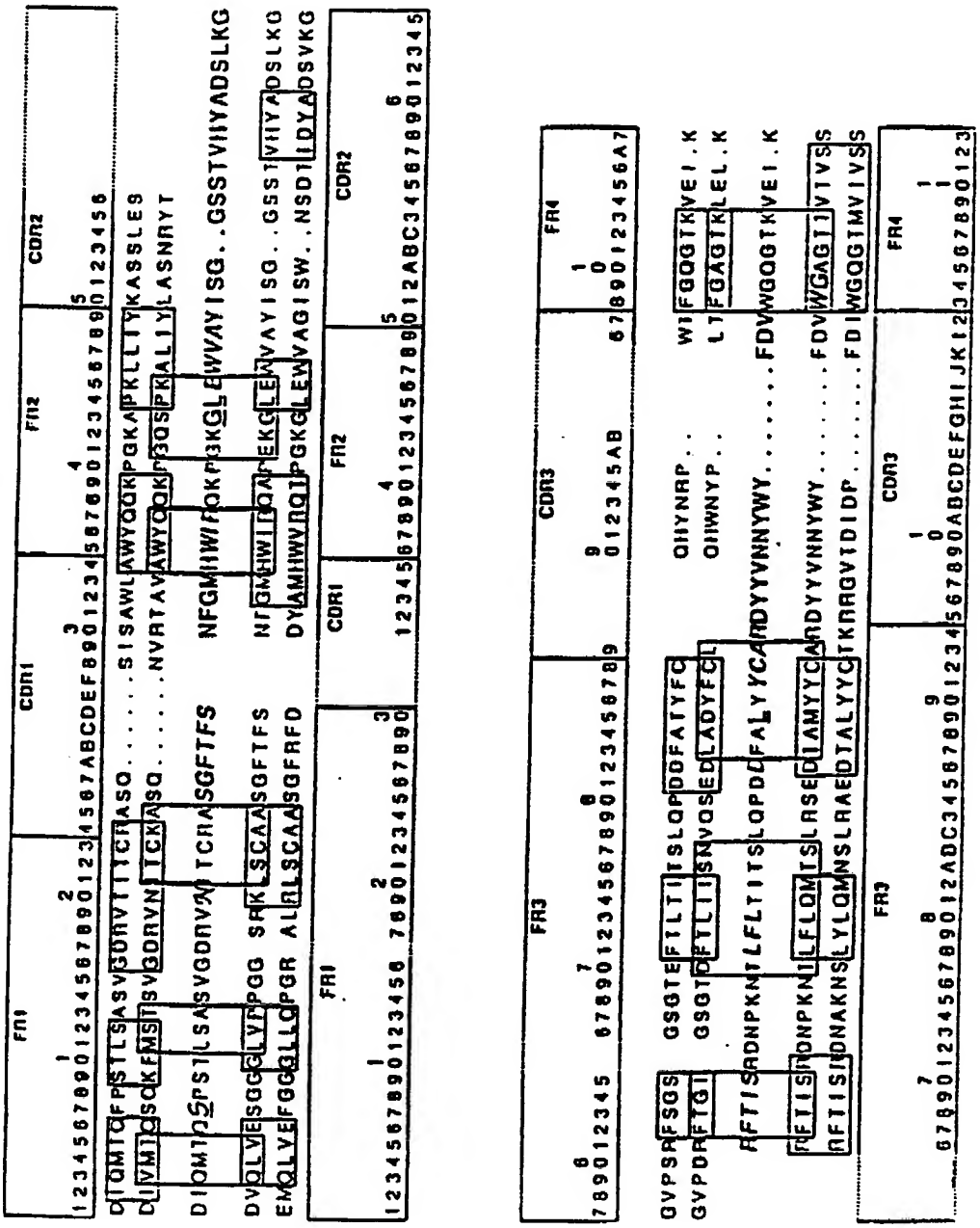
Alignment of IM9 heavy chain variable domain sequence with heavy chain variable domain sequences from immunoglobulins with known three-dimensional structure, showing SCR's and chain association residues.



Chain association residues identified in the known structure (FDL), selected as the template for IM9 chain association, are italicized bold. Residues identified in the IM9 Fv model as chain association residues are bold in the IM9 sequence.

FIG. 12

The heavy body, HB, sequence containing the light chain variable region of IM9 grafted with the Kabat-defined CDR's from the heavy chain of ZCE025, aligned with heavy and light chain sequences of ZCE025 and IM9.



IM9 ,Human Kappa Chain with ZCE-025 Heavy Chain CDR's

diqmtqfpstlsasvgdrvNitcraSGFTFS NFGMH wIRqkpgkGLkWVAy
@ ****** β** $$ ***

Framework 1 — CDR 1 — Framework 2

ISGGSSTVHYADSLKG rfTIsRDNPKNeLFltitslqpddfaMyYcAR
ββ**β******** **β $ β* β**

CDR 2 — Framework 3

DYYVNNYWYFDV Wgqgtkveikr
*

CDR 3 — -Framework 4-

Legend: lower case — IM9-human kappa-1
upper case — ZCE-025 murine gamma-1
@ — Glycosylation site
* — Framework changes supporting cdr's
$ — Domain association
β — Common to both ZCE heavy and IM9 light chain

FIG. 13

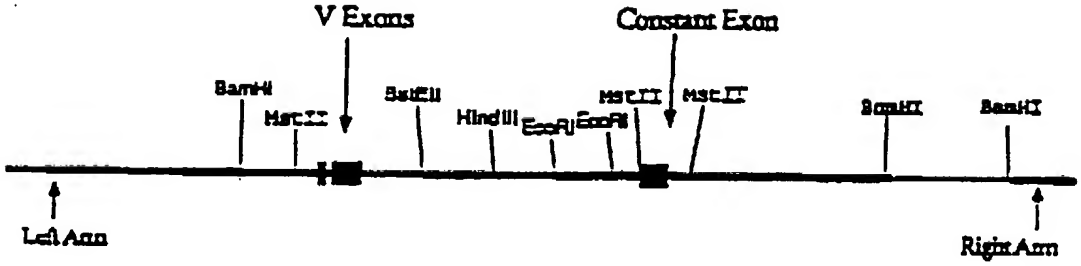

V Exons
Constant Exon
BamHI
MstII
BstEII
HindIII
EcoRI
EcoRI
MstII
MstII
BamHI
BamHI
Left Arm
Right Arm

FIG. 14

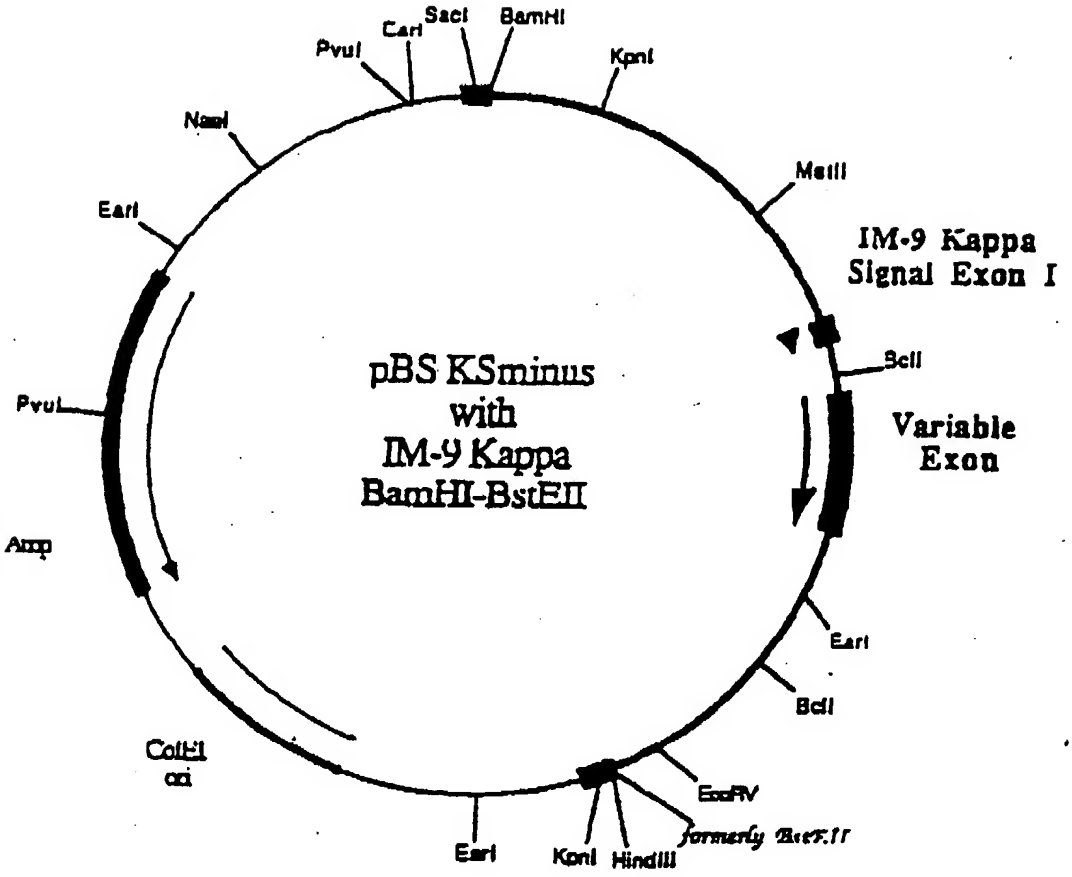

pBS KSminus
with
IM-9 Kappa
BamHI-BstEII
IM-9 Kappa
Signal Exon I
Variable
Exon
SacI
BamHI
KpnI
MaeII
BclI
EarI
EcoRV
formerly BstEII
HindIII
KonI
EarI
ColE1
ori
Amp
PvuI
NaeI
EarI
PvuI
ClaI

FIG. 15

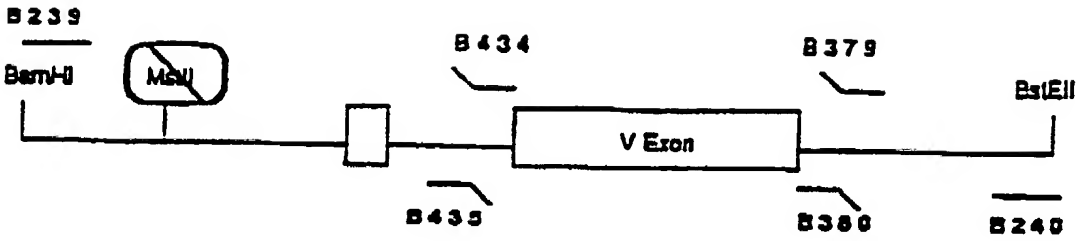


FIG. 16

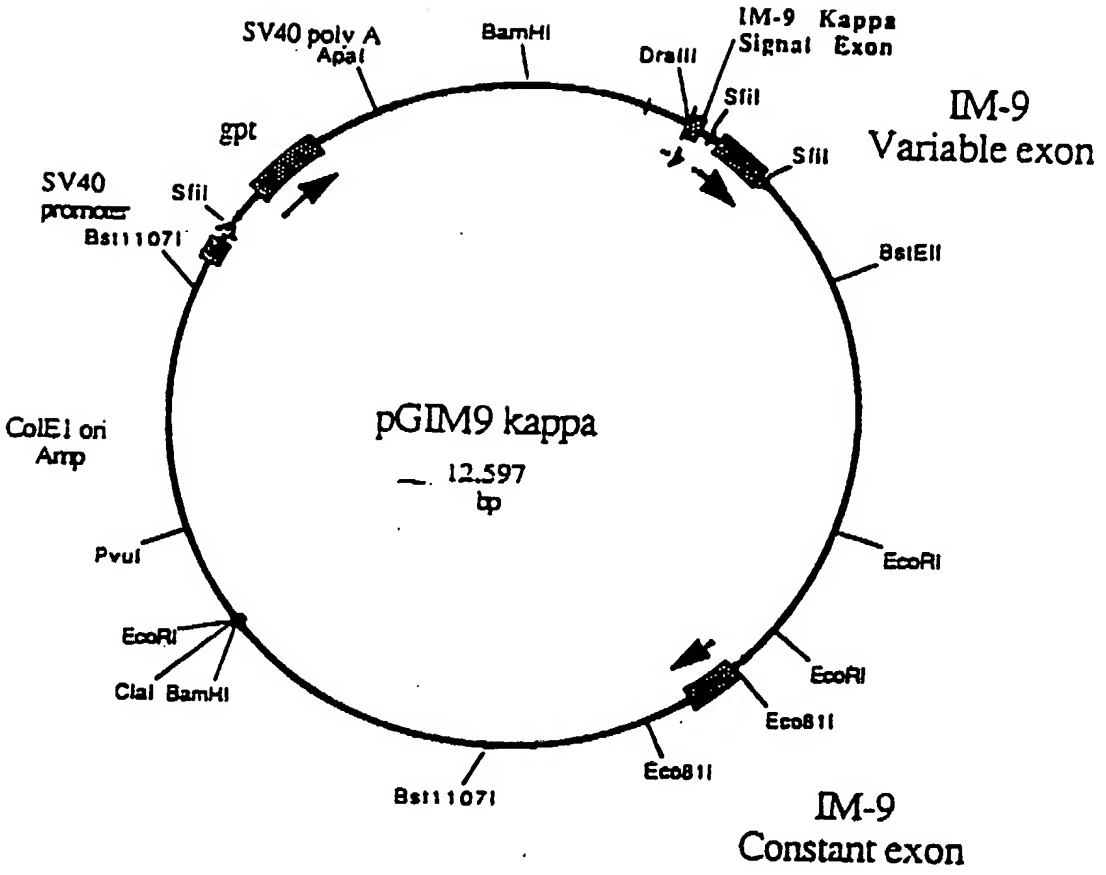

SV40 poly A
ApaI
BamHI
IM-9 Kappa
Signal Exon
DraIII
SfiI
IM-9
Variable exon
gpt
SV40
promoter
SfiI
Bst11071
BstEII
pGIM9 kappa
12,597
bp
ColE1 ori
Amp
PvuI
EcoRI
EcoRI
ClaI
BamHI
EcoRI
Eco81I
Eco81I
Bst11071
IM-9
Constant exon

FIG. 17

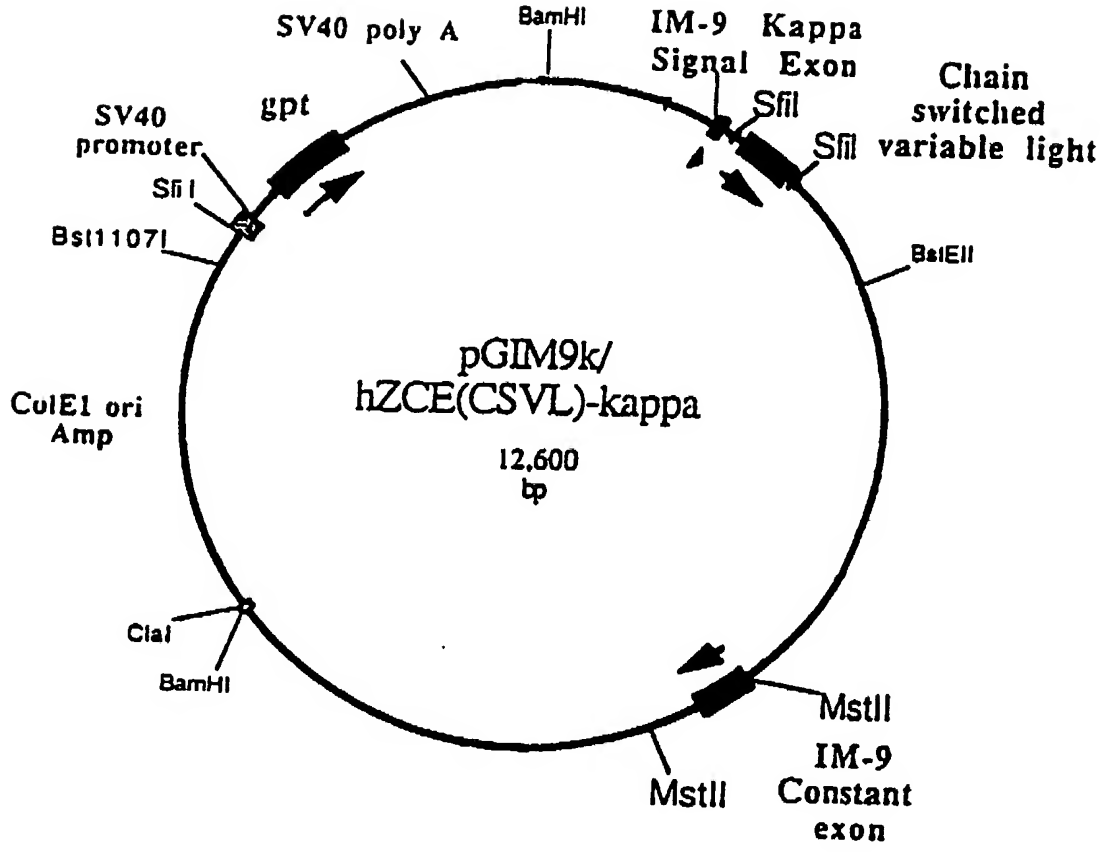

SV40 poly A
BamHI
IM-9 Kappa Signal Exon
Chain switched variable light
SV40 promoter
gpt
SfiI
SfiI
Bsl1107I
BsiEII
pGIM9k/
hZCE(CSVL)-kappa
12,600 bp
ColE1 ori Amp
ClaI
BamHI
MstII
IM-9 Constant exon
MstII

FIG. 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/10791

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 35/16; 39/00; 39/395
US CL :530/387.2; 536/23.53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/387.2; 536/23.53

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,013,653 (HUSTON ET AL.) 07 May 1991, see entire reference.	1-98
Y	Proc. Natl. Acad. Sci. Vol. 78, NO. 12, issued December 1981, Schreiber et al. "Monoclonal antibodies against receptor for epidermal growth factor induced early and delayed effects of epidermal growth factor, pages 7535-7539, see entire reference.	1-98
Y	EP, A, 0 239 400 (WINTER) 30 September 1987, see entire document.	1-98
Y	Proc. Natl. Acad. Sci. Vol. 86, issued December 1989, QUEEN ET AL., "A humanized antibody that binds to the interleukin 2 receptor", pages 10029-10033, see entire reference.	1-98

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 OCTOBER 1995	27 NOV 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DONALD E. ADAMS, PH.D. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/10791

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NATURE, Vol. 332, issued 24 March 1988, Riechmann et al. "Reshaping human antibodies for therapy", pages 323-327, see entire document.	1-98